UCRL-ID-147449

White Paper on Institutional Capability Computing Requirements

prepared by

Lynn Kissel
Michel G. McCoy
Mark K. Seager

Integrated Computing and Communications Department (ICCD)
Computation Directorate
Lawrence Livermore National Laboratory

and the open computing scientists at LLNL

for the

Institutional Computing Executive Group (ICEG)

Associate Director for Computation

Deputy Director for Science and Technology

January 29, 2002

This work was performed under the auspices of the U.S. Department of Energy by the University of California, Lawrence Livermore National Laboratory under Contract No. W-7405-Eng-48.

Contents

Executive Summary

This paper documents the need for a rapid, order-of-magnitude increase in the computing infrastructure provided to scientists working in the unclassified domains at Lawrence Livermore National Laboratory. This proposed increase could be viewed as a step in a broader strategy linking hardware evolution to applications development that would take LLNL unclassified computational science to a position of distinction, if not preeminence, by 2006. We believe that it is possible for LLNL institutional scientists to gain access late this year to a new system with a capacity roughly 80% to 200% that of the 12-TF/s (twelve trillion floating-point operations per second) ASCI White system for a cost that is an order of magnitude lower than the White system.[1] This platform could be used for first-class science-of-scale computing and for the development of aggressive, strategically chosen applications that can challenge the near PF/s (petaflop/s, a thousand trillion floating-point operations per second) scale systems ASCI is working to bring to the LLNL unclassified environment in 2005.

As the distilled scientific requirements data presented in this document indicate, great computational science is being done at LLNL—the breadth of accomplishment is amazing. The computational efforts make it clear what a unique national treasure this Laboratory has become. While the projects cover a wide and varied application space, they share three elements—they represent truly great science, they have broad impact on the Laboratory's major technical programs, and they depend critically on big computers.

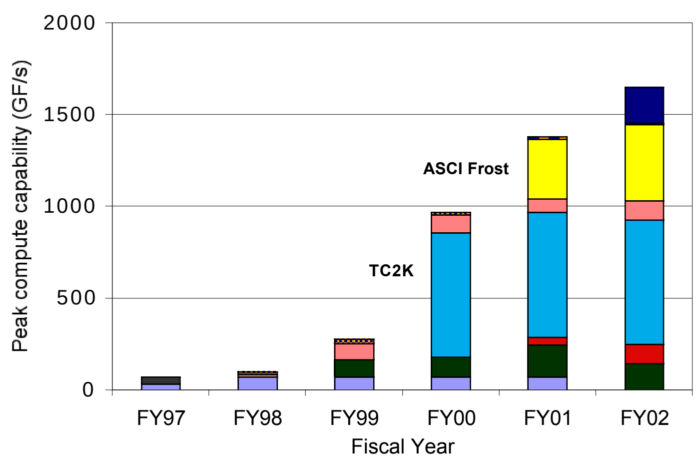

Figure 1. Growth in the capacity of M&IC is shown for FY97–FYF02. The total capacity, currently about 1.6 TF/s, is dominated by the TC2K and ASCI Frost systems.

It is the third of these three elements that concerns us here. While many of the scientific applications are first-of-class in their fields, their quality is not assured without significant and sustained investments in our institutional computing infrastructure. There is growing demand for dedicated, powerful parallel computational systems, and this demand is increasingly unsatisfied. Today, the growth of leading-edge programs at this Laboratory is constrained by access to computational resources. At the heart of the matter is the fact that SSP hardware on our floors cannot be dominated by institutional use, and the institutional hardware is being left behind relative to rapid buildup at other institutions. Currently, Multiprogrammatic and Institutional Computing (M&IC) scientists have access to about 1.6 TF/s sustained (see Fig. 1), across several systems, both ASCI and institutional.

[1] How the proposed architecture compares with White depends on the scientific application chosen for comparison. A suite of tests on many codes run on the SSP Linux system provides the basis for the projection of 80% to 200%.

Analyzing the data presented to us in an M&IC questionnaire by 27 scientific efforts at LLNL reveals a problem that can be resolved only with an increase in capacity by a factor of 5 to 8.

- A straightforward analysis of the data suggests a need for at least 8 TF/s to put a significant dent in demand coming from all scientific quarters of LLNL and to allow existing institutional projects to move forward expeditiously.
- The institution could instead elect to make room for one strategically selected computational thrust, perhaps even a Grand Challenge. This could come from materials science, environment, or biology, for example (requiring >4 TF/s). However, the machine would also have to make room for programmatic co-investment in the hardware (>2 TF/s). This suggests a minimum requirement of 6 TF/s. *Such triage leaves many worthy projects without significant access, because there is no room for institutional applications outside of the strategic thrust.*
- If there is, in addition to the above, an effort to meet the broad requirements coming from the full spectrum of institutional arenas (>4 TF/s), the minimum requirement is for a 10 TF/s computer.

A survey of national developments can provide useful ancillary data. A procurement in the 6–10 TF/s range would put LLNL approximately equivalent to unclassified-computing leaders, such as LBNL's National Energy Research Scientific Computing Center, the NSF's Pittsburgh Supercomputer Center, and soon PNNL and the NSF Distributed Terascale Facility (DTF) at 11 TF/s (see Fig. 2).

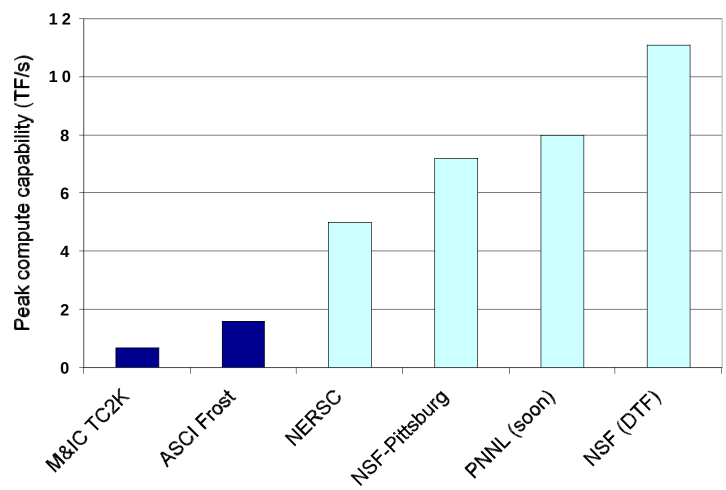

Figure 2. The capacity of the largest open-partition computing resources at the Laboratory (TC2K and Frost, shown in dark blue) and those of other institutions (light blue). LLNL currently has systems of the order 1 TF/s; DOE-SC and NSF have systems of the order 10 TF/s.

We propose three related steps as part of a broad strategy to bring LLNL unclassified computing to preeminence by 2006 (see Fig. 3).

1. **We request support for a procurement in FY02 to put capability on the floor to carry the institution into FY04 at the 5.5–11 TF/s level**. For the smaller system, the institution would need to make strategic decisions regarding which efforts will see ample resources and which will not. For the larger system, it will be possible to accommodate (1) a focused institutional strategic challenge, (2) programmatic investment allocations, and (3) the broad spectrum of institutional users requiring additional access levels outside of the scope of the strategic challenge. A mid-range system at the 8 TF/s level is also possible and would serve all three purposes, albeit with trade-offs.
2. **We suggest that the LLNL science community work to put together at least one comprehensive strategic initiative for the consideration of the institution**. The responses to the M&IC questionnaire show that there is a critical mass of first-tier applications codes

and scientific expertise to be harnessed for any number of strategic institutional computational investments. This could take the form of a Grand Challenge. A first-class computing infrastructure in the open should be associated with an institutionally sanctioned science thrust.

3. **We hope the institution will continue to track the progress of the ASCI Program in bringing to LLNL an IBM advanced-architecture 0.2–0.4 PF/s system in late FY04 or early FY05.** An institutional investment in the hardware would be called for then that could give LLNL a substantial ownership position in the computer. The strategic initiative efforts mentioned in item 2 above should be targeted in part at this system. These could be developed on the Linux cluster proposed for FY02. Such a combined applications development effort and an innovative architectural step would put LLNL into a distinctive science and technology position.

M&IC has hit a speed bump and is at a crossroad. Without a significant increase in capability, it will devolve into a capacity operation with no ability to enable science of scale. With an increase as is requested here, it could leverage the best of the opportunities made available by ASCI to create opportunity for all LLNL scientists. The path we can follow is not along a purgatorial and mined-out technology curve, providing a meager factor-of-2 increase every couple of years. Instead we see the possibility to leapfrog across—or even to straddle—technology curves. We can take advantage of the steep part of their slopes for rapid progress and the flatter parts for

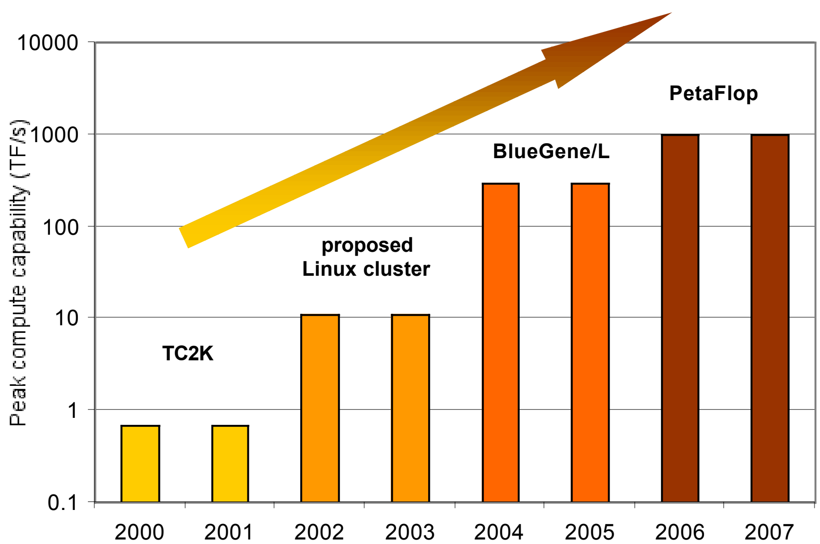

Figure 3. A possible platform path for LLNL preeminence.

stability (see Fig. 4). Soon, we could straddle three technologies simultaneously: computers based on (1) workstation or mainframe class processors with vendor-supplied operating systems and software, (2) IA-32 or IA-64 processors with Linux and Open Source software, and (3) cell-based systems from IBM. This advantage is currently unique at this magnitude to LLNL because of our relationship with IBM, our historical development on Linux systems, and our current heavy integration work at Livermore Computing on Linux-based architectures at the high end.

Section 1 presents more details on the challenging situation faced by unclassified computing at LLNL. Section 2 gives an analysis of the input provided by the 27 science teams represented in this document and an estimate of the capacity required. Section 3 provides our recommendations.

The appendices provide extensive information on the science requirements responses (Appendices A–F), a detailed technical description of the proposed Linux solution (Appendix G), some resource allocation models for running the machine for ICEG to consider (Appendix H), and a description of the architecture and promise of BlueGene/L (Appendix I).

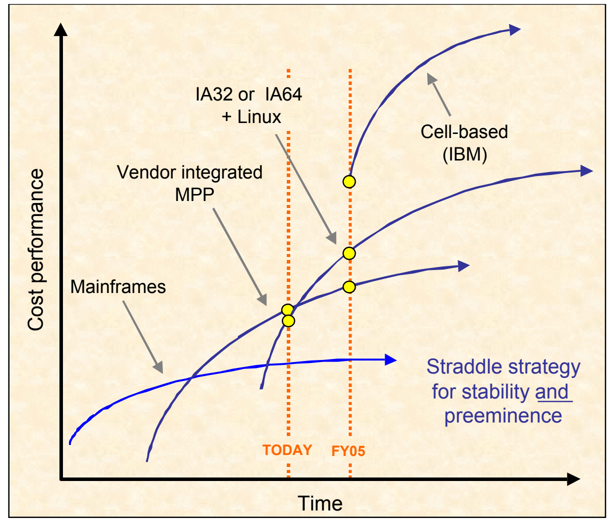

Figure 4. M&IC is moving across technology curves that provide increasingly cost effective computing capability while maintaining stability.

There is also a separate addendum, "Unclassified Computing Capability: User Responses to a Multiprogrammatic and Institutional Computing Questionnaire," that provides the compiled input from the contributing science teams. This document has collected the requirements of LLNL's scientists, describing their accomplishments to date under M&IC, their impact on programs and the institution, the number of FTEs focused on this work, and the current limitations they face from paucity of cycles. This material was used as the basis for the analysis and recommendations presented in this document.

1. Status of Open-Partition High-Performance Computing

In the latter half of the 1990s, the maturation of parallel computing technology made it possible to accelerate the development of production-level 3D scientific applications requiring super-teraflop computational capability. Indeed, the Stockpile Stewardship Program proposed and followed through with ASCI. LLNL confronted the reality that one of its major programs was embarking on an adventure that had the potential to revolutionize how science would be done in the next century.

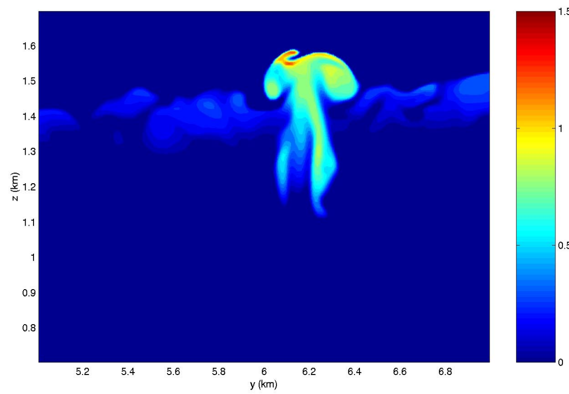

Figure 1.1. The formation of cumulus clouds in the atmosphere as simulated in a science run on ASCI White. Understanding cloud formation is an important component of accurate weather and climate simulations.

At this juncture, the Director's Office made the determination that a mechanism had to be found to leverage this initiative so that all scientific areas at the Laboratory could follow in the channel cut through the ice by ASCI. LLNL scientists, in particular those engaged in unclassified research, were competing with science teams from across the country for funding, access to experimental apparatus, access to high-performance computing infrastructure, and, ultimately, for results and recognition. It was strategic that LLNL would seek to support its scientists, regardless of programmatic connection, in this arena. (Important unclassified simulations like that shown in Fig. 1.1 by David Stevens have limited access to ASCI platforms, such as during science runs that are used to debug and stabilize new platforms.)

From this notion was born Multiprogrammatic and Institutional Computing (M&IC). M&IC is truly institutional. Many directorates invest, and the institution invests. The board of directors is composed of distinguished computational scientists from across LLNL. The growth of M&IC since 1997 has been significant (see Fig. 1 in the Executive Summary); the total capacity currently available to M&IC scientists is about 1.6 TF/s.

Two related factors have emerged in 2002 that threaten M&IC's continued success. The first is that in FY01 a number of LLNL teams requested significantly larger levels of access. Some teams are dropping projects for lack of access. A check of the queues for ASCI Frost and TC2K on Friday evening, January 25, showed more than 100 jobs waiting in the queues; some had been waiting for several days. This kind of "success" limits progress.

The second factor is that M&IC capability systems are no longer competitive with the best unclassified systems outside of LLNL. This was highlighted in the Executive Summary. This represents a reversal, because in the past LLNL institutional scientists enjoyed the equivalent of the best systems available on the outside, and now the Laboratory will soon be behind by *at least* an order of magnitude. This has left our scientists at a disadvantage with respect to those scientists who qualified for significant access to those sites.

In responding to this challenge, LLNL is not without its strengths. The first is that the leverage from ASCI continues to work its magic. ASCI has provided the R&D funds that make it possible for M&IC to invest in high-performance but low-cost infrastructure systems, usually before most other sites can do so. Two such examples are discussed in the appendices to this paper—Linux clusters in 2002 and 2003 (Appendix G) and the IBM BlueGene/L in 2004 and 2005 (Appendix I). A Linux cluster (Fig. 1.2) will soon be in classified production for the SSP. Second, again because of ASCI investments, LLNL has the benefit of one of the most experienced and well-staffed scientific computing centers in the world. An investment in hardware is leveraged by attention from experienced integrators, operators, and services staff and from a well-engineered

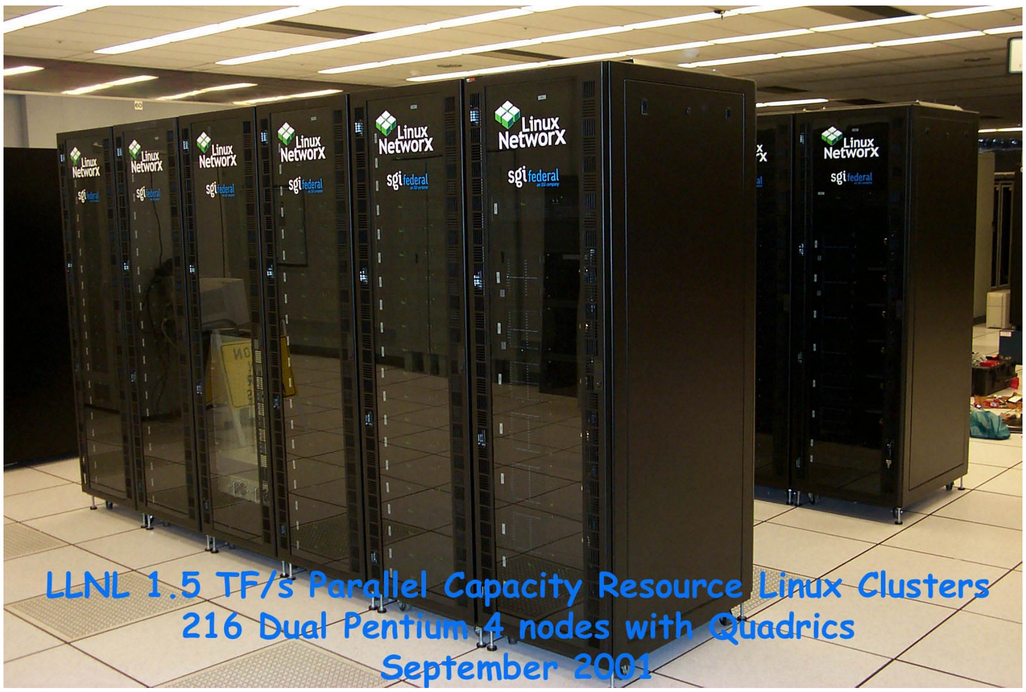

Figure 1.2. The Parallel Capacity Resource is demonstrating that IA-32–based Linux clusters can provide cost-effective computing resources for the Laboratory.

foundation in networks and storage. All of this mitigates considerably the risks inherent in investing in the newest and best cost performance technologies.

2. Requirements Summary

In September 2001, 27 users of M&IC resources described their near-term computing requirements by responding to a 12-question survey. Appendices A–F list the questions and summarize the responses; the detailed responses are available as a separate document.[2] In this section we interpret the requirements that were gleaned from the 27 responses.

Table 2.1. The 27 computationally intensive projects that responded to the request to document their computational needs are listed with an arbitrarily assigned number and Project ID for reference.

	Project ID	Submitted by	Title	FTEs
1	ALPS	Milo Dorr	ALPS (Adaptive Laser Plasma Simulator) Project	3.15
2	DJEHUTY	Dave Dearborn, Don Dossa	The Djehuty Project	2
3	AMRh	Allen Kuhl, Jeff Greenough	AMRh Code Development Project	4.5
4	Fermion MC	Malvin Kalos	Physics Problems in Very High Dimensions	2
5	DD-ICF	Steven Langer	Direct-Drive ICF	1
6	Z3	Barbara Lasinski, Bruce Langdon, Bert Still	Z3 Project	1
7	Mat-Shock	Brian Wirth, James Stolken, Maria Caturla	Shock Propagation in Materials	2.5
8	Mat-Rad	Tomas Diaz de la Rubia, Brian Wirth, Bill Wolfer, Maria Caturla	Radiation Damage of Materials	3.5
9	Cell Modeling	Carl Melius	Modeling of Biological Cells for Bioterrorism and Health	3
10	Biochem	Andrew Quong, Chris Mundy	First-Principles Molecular Dynamics for Understanding Fundamental Biochemical Interactions	3
11	CompBio	Mike Colvin, Felice Lightstone	Computational Biology Project	4
12	GFMD	Lin Yang	GFMD (Greens Function Molecular Dynamics) Project	3
13	BOUT	Xueqiao Xu	BOUT (Boundary-plasma Turbulence) Project	2
14	NuclStruct	Erich Ormand	Ab Initio Nuclear Structure	2
15	JEEP	Giulia Galli, Francois Gygi	The Jeep and Quantum Simulations Projects	16
16	PHENIX/HBT	Ron Soltz	PHENIX/HBT	11
17	MD3D	Jim Belak, Robert Rudd	Microscopic Origins of Dynamic Fracture Project	3
18	pF3d	Bert Still, Ed Williams, Richard Berger, Bruce Langdon, Laurent Divol	pF3d - Predictive Laser-Plasma Interaction Modeling	4
19	NIF gas	Steven Sutton	Gas Distortion Characterization for the NIF	2.8
20	EIGER codes	Rob Sharpe	Electromagnetic Effects in High-Frequency ICs	3
21	NDE	Pat Roberson	Nondestructive Characterization	30
22	E3D	Shawn Larsen	Seismic and Acoustic Wave Propagation	10
23	HP-CFD	Bob Lee, Don Ermak, Stevens Chan	High-Performance CFD Models	3
24	NUFT-C	Bill Glassley, John Nitao	NUFT-C Project	10
25	HR-GCS	Phil Duffy	High-Resolution Global Climate Simulations	1
26	AtmosChem	Doug Rotman, Cyndi Atherton, Peter Connell, Cathy Chuang, Jane Dignon, Dan Bergmann, John Tannahill, Philip Cameron-Smith	Global Atmospheric Chemistry	12
27	Earthquake	Charles Noble	Morrow Point Dam Earthquake Analysis	1

2.1 General Observations

There are many important scientific projects at this Laboratory with applications that depend critically on access to very high performance computing. We draw several important general conclusions from the responses.

1. An enormous role is being played in **experimental science** by these computational projects. Almost all of the projects are directly involved with the use of experimental results or involve direct comparison with experiment. Many of the projects are involved in the design of experiments, while some are used for the design of facilities. Nearly half of the projects have some connection to NIF, including one that focuses on understanding the thermally

[2] "Unclassified Computing Capability: User Responses to a Multiprogrammatic and Institutional Computing Questionnaire," available from the Integrated Computing and Communications Department, LLNL, Jan. 29, 2002.

driven optical distortions in order to increase the frequency of shots. To be truly effective, these efforts must rely on a generation of 3D simulations, using codes like ALE3D. Commercial codes (such as FIDAP) are generally designed for smaller problems and target shared-memory systems that are not focused on technology-of-scale requirements.

Many projects have achieved such sophistication that direct comparisons between full-scale experiment and simulations based on ab initio models are now possible but are often limited by the available computing resources. For example, the ALPS Project is ready to perform 3D runs at the size of the experimental configuration currently being used on the Omega laser facility at the University of Rochester in support of the NIF Program. As another example, the AMRh Project is ready to perform 3D simulations that would more fully resolve the turbulent breakup of vortex rings produced by shock/sphere interactions that are under experimental investigation (RIKSPHERE Omega Laser Experiments). Finally, the pF3d Project is ready to perform the first simulation on the entire Nova beam volume and perhaps even the volume of a NIF outer beam.

Table 2.2. The projects span the Laboratory's technical directorates.

	Project ID	Technical Directorates Involved in this Project								
		BBRP	C&MS	Comp	DNT	EED	Eng	NIF	NAI	PAT
1	ALPS			team organization	involved organization					involved organization
2	DJEHUTY			team organization						team organization
3	AMRh				team organization					
4	Fermion MC			team organization	team organization					team organization
5	DD-ICF				team organization					
6	Z3				team organization					
7	Mat-Shock		team organization		involved organization		involved organization	involved organization		involved organization
8	Mat-Rad		team organization				involved organization			involved organization
9	Cell Modeling		team organization							
10	Biochem		team organization							
11	CompBio	team organization		involved organization						involved organization
12	GFMD		involved organization							team organization
13	BOUT									team organization
14	NuclStruct									team organization
15	JEEP	involved organization		team organization					involved organization	team organization
16	PHENIX/HBT		involved organization				involved organization			team organization
17	MD3D			involved organization						team organization
18	pF3d				team organization			involved organization		team organization
19	NIF gas						team organization	involved organization		
20	EIGER codes						team organization			
21	NDE						team organization			
22	E3D			team organization		team organization	team organization	involved organization	involved organization	involved organization
23	HP-CFD					team organization			involved organization	
24	NUFT-C					team organization				
25	HR-GCS			involved organization		team organization				
26	AtmosChem					team organization				
27	Earthquake						team organization			

team organization   involved organization

2. Computation is now a mainstream method in **theoretical science** at LLNL. Computation is essential at the level at which highly simplified but analytically intractable models are explored or complex multiphysics phenomena must be understood quantitatively. As we understand more and more truly basic science, the Laboratory is looking to computation to make the vital quantitative connections among disparate phenomena that constitute the foundation of both pure and applied science.

   For example, the Z3 Project is aimed at development of a state-of-the-art, first-principles predictive capability for laser–plasma interactions, critically important to the ICF Program as well as for other programs at the Laboratory. As a second example, the Mat-Shock Project is obtaining a basic understanding of the interaction of shocks with material microstructure, a subject critically important to many programs at the Laboratory.

3. The total **computing needs** of these projects far exceed the current M&IC capability and capacity. The pace of progress on many of these projects is being hindered by lack of capacity, and access to additional computing resources will result in faster progress.

   For example, the JEEP and Quantum Simulations Projects, under the direction of G. Galli, have put a number of subprojects on hold, based on the paucity of cycles. This team currently receives among the largest allocations on both TC2K and ASCI Frost. Only an order-of-magnitude increase in cycles will result in a significant improvement in the challenges faced here. This team has resorted to looking elsewhere for additional cycles because the institutional pool is insufficient—a warning sign for M&IC. In short, for a number of efforts, further progress is not possible without a considerable expansion to more capable computing resources.

Table 2.3. Most of these computationally intensive projects have a close connection with experiment.

	Project ID	Connect to exp?	Connect to NIF?	Exp design?	Facility design?
1	ALPS	yes	yes		
2	DJEHUTY				
3	AMRh	yes	yes		
4	Fermion MC				
5	DD-ICF	yes	yes		
6	Z3	yes	yes		
7	Mat-Shock	yes	yes		
8	Mat-Rad	yes	yes		
9	Cell Modeling				
10	FP-Biochem	yes			
11	CompBio	yes		yes	
12	GFMD	yes			
13	BOUT	yes			
14	NuclStruct				
15	JEEP	yes	yes		
16	PHENIX/HBT	yes			
17	MD3D	yes			
18	pF3d	yes	yes	yes	yes
19	NIF gas	yes	yes		yes
20	EIGER	yes			
21	NDE	yes	yes		
22	E3D	yes	yes	yes	yes
23	HP-CFD	yes	yes	yes	
24	NUFT-C	yes		yes	
25	HR-GCS	yes			
26	AtmosChem	yes		yes	
27	Earthquake	yes		yes	yes

Table 2.4. Many of the projects share common approaches, but they differ markedly in the details of the implementations.

	Project ID	MD	AMR	LPI	CFD	hydro	rad-hydro	rad trans
1	ALPS		x	x		x	x	
2	DJEHUTY					x	x	
3	AMRh		x			x	x	
4	Fermion MC							
5	DD-ICF			x		x	x	
6	Z3			x				
7	Mat-Shock	x						
8	Mat-Rad	x						
9	Cell Modeling							
10	FP-Biochem	x						
11	CompBio	x						
12	GFMD	x						
13	BOUT					x		
14	NuclStruct							
15	JEEP	x						
16	PHENIX/HBT							
17	MD3D	x						
18	pF3d			x				
19	NIF gas					x		
20	EIGER							
21	NDE							
22	E3D							
23	HP-CFD				x			
24	NUFT-C							
25	HR-GCS				x			x
26	AtmosChem							
27	Earthquake							

LLNL is now in the forefront of the evolution toward effective and practical computational science in all its forms. To continue in that role, we must continue to provide a wide community—not just the high and conspicuous end—with the tools they need to advance their scientific research.

One common misconception about computational science is that large machines are needed only for production, while small machines suffice for program development. Although this is true in some cases, especially where well-known and tested methods are applied to larger problems than before, it is not true when paradigms change as the challenges change. Going from one to two to three dimensions in studying fluid dynamics was certainly not a mere change of the scale of the computations. It would have been impossible to use a UNIVAC I to prepare for modern computational studies of turbulence. Similarly, in the study of many-body physics, the challenges are quite different in treating classical versus quantum systems and in passing from a few to hundreds or millions of particles.

Table 2.5. One important programmatic impact is listed for each project.

Project ID		Programmatic Impact of Project
1	ALPS	Predictive modeling of laser plasma interaction (LPI), fundamental for the design and analysis of laser-driven fusion experiments.
2	DJEHUTY	Improved determination of the size, age, and composition of the universe through 3D modeling of stars.
3	AMRh	Accurate calculation of compressible, high-Reynolds-number flows--critical to understand turbulence and mixing.
4	Fermion MC	Fundamental understanding of opacities and equations of state from first principles.
5	DD-ICF	Predictive capability to simulate direct-drive ICF.
6	Z3	State-of-the-art, first-principles predictive capability for LPI.
7	Mat-Shock	Basic understanding of interaction of shocks with material microstructure.
8	Mat-Rad	Predict changes in microstructure and properties of materials exposed to radiation.
9	Cell Modeling	Understanding the functioning of microbial pathogens at the cellular level.
10	FP-Biochem	Model molecular targeting, a technique for engineering antibodies to detect and ultimately kill toxic cells.
11	CompBio	Develop a laboratory core competency in computational biology. First-principles molecular dynamics simulations for realistic biological systems.
12	GFMD	Quantum-based atomistic simulations of materials properties in transition metals, for multiscale modeling of strength and failure.
13	BOUT Proj	Self-consistent modeling of plasma and neutral-particle transport in the edge plasma of magnetic fusion devices.
14	NuclStruct	First-principles description of the structure of light nuclei.
15	JEEP	Predict physical and chemical properties of matter with great accuracy, using advanced quantum simulation techniques.
16	PHENIX/HBT	Detect and characterize the quantum chromodynamics phase transition (the melting of protons and neutrons into a plasma of their constituent quarks and gluons).
17	MD3D	Model the nucleation and growth of voids in ductile metals during dynamic fracture.
18	pF3d	Develop a predictive capability for LPI.
19	NIF gas	Detailed simulations of the flow field in trapped gas volumes on NIF.
20	EIGER	Modeling of electromagnetics effects in high-frequency integrated circuits.
21	NDE	Nondestructive characterization of objects and materials at resolution and detail not previously possible.
22	E3D	Model and characterize seismic and acoustic wave propagation in the earth and other material.
23	HP-CFD	High-performance computational fluid dynamics models for simulating flow and dispersion of hazardous materials over urban areas.
24	NUFT-C	Rigorously account for the coupled physical and chemical processes that occur as subsurface water migration takes
25	HR-GCS	Present and future global climates at higher spatial resolution than has ever been used.
26	AtmosChem	Understand the effects of natural and anthropogenic activities on the distribution of important atmospheric chemical
27	Earthquake	Influence how the U.S. determines dam safety risks.

In developing radically new methods for new areas, it is therefore necessary to have computers of high enough capability to allow rapid experimentation with methods while attacking realistically large systems. Such experimentation should be completed within a few hours (or at most days), rather than weeks or months, to permit progress at a reasonable pace. For this reason, it is often extremely difficult to make credible estimates of the computational resources needed for particular advances. An example of such a situation is that faced by the ocean and atmospheric codes (see submissions by Duffy [Project ID 25], Fig. 2.1 and Rotman [Project ID 26], Fig. 2.2), in which distinct challenges are faced at each level of resolution selected. Consequently, development work needs to employ essentially the same scale of resources as production calculations.

2.2 Strategic Challenges—Ab Initio Modeling and Materials Science

About a quarter of the responses (seven) involved some form of simulation based on molecular dynamics (MD). Physics-based, *experimentally validated* simulation of materials is a significant challenge that could revolutionize the way we approach all kinds of systems, from the stockpile (certification and remanufacturing), to future environmentally benign sources of energy, to biology and bioterrorism. Validated MD simulations are critical to material discovery (i.e., the design, process and manufacturing, and performance prediction of materials).

Ab initio modeling with access to sufficient computing resources is both big science and critical to LLNL's future in several areas, such as materials and biology. Both the machines and the codes are unique LLNL capabilities, worthy of continued investment, with the return on investment being the unique science of scale that we can do. Predicting material lifetimes, understanding and curing diseases, detecting and combating biological and chemical terrorism, explaining the behavior of stars, designing the next generation of electronic components (integrated circuits), analyzing earthquake hazard, understanding our climate and human influence on the environment are all scientific impacts that will result from successful implementation of these projects.

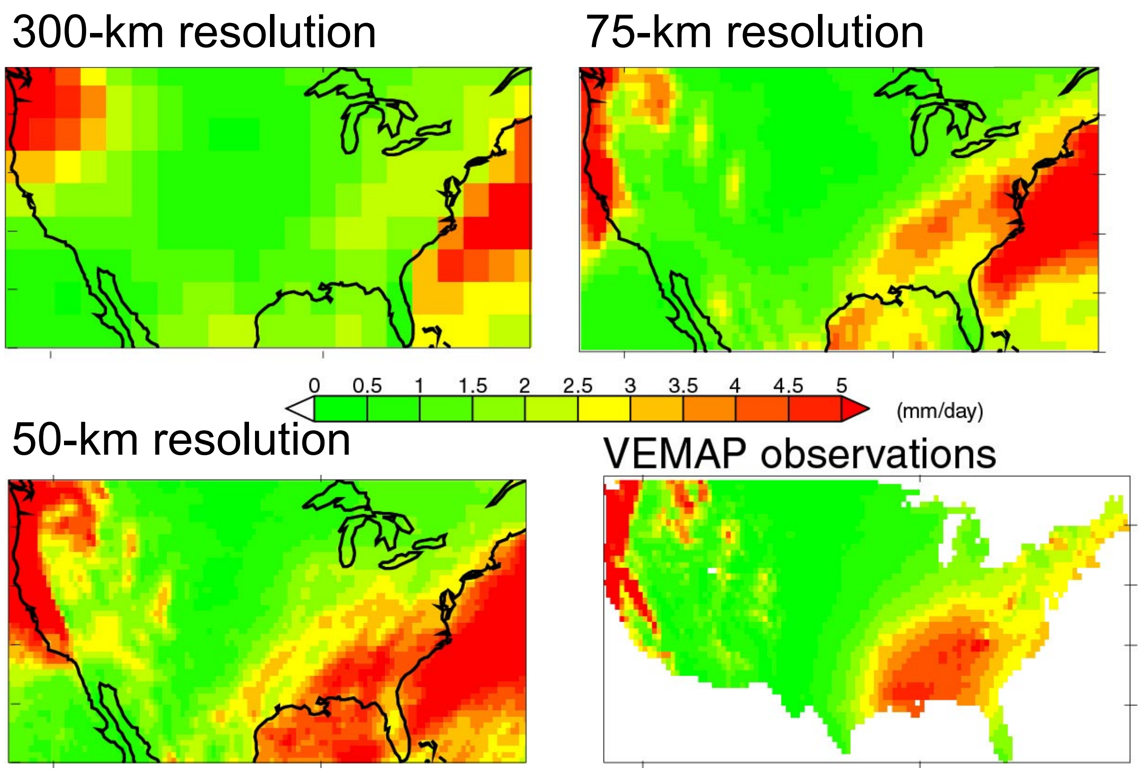

Figure 2.1 Model simulations of winter precipitation at three different resolutions and observations of the same quantity on a 50-km grid, demonstrating the importance of resolution in the fidelity of simulations. The number of computations required for 3D simulations scales as the cube of the linear resolution, taxing the capability of our largest computers.

In Section 2.3, we normalize the aggregate requirements coming from the institution relative to a nominal 8 TF/s system. An examination of Table 2.6 reveals that together just *two* well-known and respected codes, one coming from Chemistry (MDCASK [Project ID 8]) and the other from PAT/Computation (JEEP, [Project ID 15], Fig. 2.3) saturate an 8 TF/s system. A significant and comprehensive institutional thrust in computational materials science would require the involvement of both of these code teams as well as participation covering other necessary pieces of the physical spectrum (mesoscale and continuum; see Fig. 2.4[3]). Such is the demand coming from just one possible, focused strategic challenge.

[3] From a presentation by Tomás Díaz de la Rubia.

Table 2.6. Sample scheduling of a new M&IC capability resource shows that the Laboratory can fully utilize a 10-TF/s system immediately.

		Sample Scheduling of an 8-TF/s Resource, in System Weeks											
	Project ID	CY02 Q1	CY02 Q2	CY02 Q3	CY02 Q4	CY03 Q1	CY03 Q2	CY03 Q3	CY03 Q4	CY04 Q1	CY04 Q2	CY04 Q3	CY04 Q4
1	ALPS		0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5
2	DJEHUTY		2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0
3	AMRh	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
4	Fermion MC		3.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0
5	DD-ICF							2.0	2.0	2.0	2.0	2.0	2.0
6	Z3		1.0	4.0	4.0	4.0	4.0	4.0	4.0	4.0	4.0	4.0	4.0
7	Mat-Shock	1.2	1.2	1.2	1.2	1.2	1.2	1.2	1.2	1.2	1.2	1.2	1.2
8	Mat-Rad	4.0	4.0	4.0	4.0	4.0	4.0	4.0	4.0	4.0	4.0	4.0	4.0
9	Cell Modeling				0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5
10	FP-Biochem				1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
11	CompBio	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0
12	GFMD			3.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0
13	BOUT		1.5	1.5	1.5	1.5	1.5	1.5	1.5	1.5	1.5	1.5	1.5
14	NuclStruct	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5
15	JEEP	4.0	4.0	4.0	4.0	4.0	4.0	4.0	4.0	4.0	4.0	4.0	4.0
16	PHENIX/HBT	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5
17	MD3D		1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
18	pF3d	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
19	NIF gas			0.2	0.2	0.2	0.2	0.2	0.2	0.2	0.2	0.2	0.2
20	EIGER				1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
21	NDE			2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0
22	E3D	3.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0
23	HP-CFD		0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5
24	NUFT-C	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5
25	HR-GCS	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5
26	AtmosChem	3.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0
27	Earthquake	3.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0
	total allocated	24.2	33.7	41.9	44.4	44.4	44.4	46.4	46.4	46.4	46.4	46.4	46.4
	% allocated	202%	281%	349%	370%	370%	370%	387%	387%	387%	387%	387%	387%

(blank)	not needed	2.0	needs more resource
0.5	good fit	2.0	needs MUCH more resource

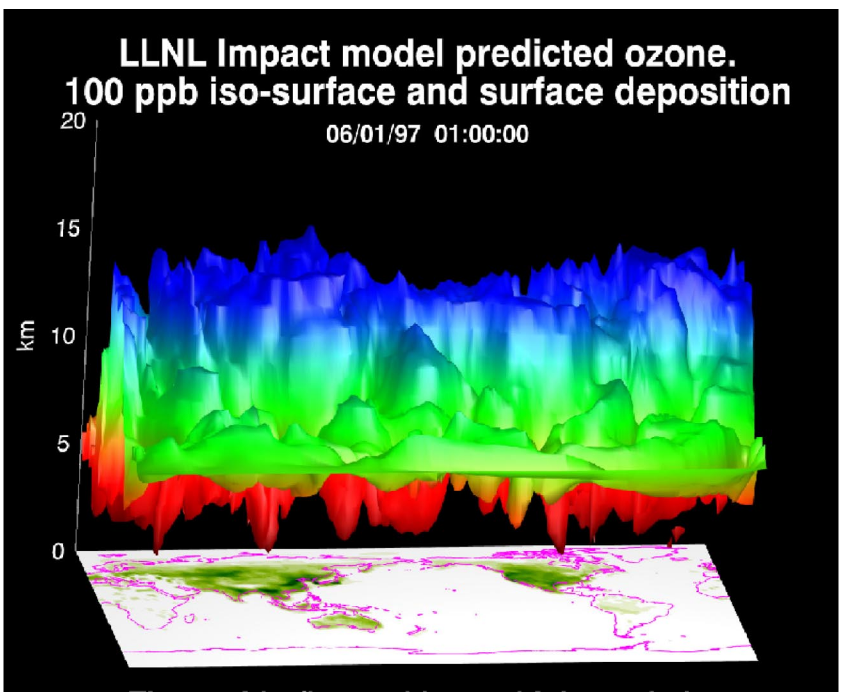

Figure 2.2. Accurate modeling of the atmosphere is needed to understand the effects of natural and anthropogenic activities on the distribution of important atmospheric chemical species.

2.3 Considerations in Sizing the Needed Resource

To understand the size of the resource that is needed to satisfy the computational requirements of these projects, we examine the sample scheduling of an 8 TF/s system. The results are displayed in Table 2.6. Some conclusions are immediately obvious.

- **Overall demand.** We are able to oversubscribe an 8 TF/s resource by a factor of 2 to 3 immediately, and easily by a factor of 4 within 2 years. This is the straightforward analysis—basically meeting the stated needs of LLNL's scientists using the existing logic: namely, important institutional efforts (as validated by the CSST) need adequate access to make progress. This concept is at the philosophical foundation of M&IC. While it is clear that estimates of computational needs provided by these projects are necessarily imprecise, we are confident that the Laboratory can fully and easily utilize a new M&IC computational resource as large as 10 TF/s immediately.
- **Impact of institutional strategic thrusts**. We discussed in Section 2.2 how a coordinated and comprehensive *materials science* effort consisting of perhaps four efforts (codes) spanning the physical scales focused on some important application of national interest might come to dominate an 8 TF/s system. An examination of Table 2.6 shows that *environmental efforts* (Project IDs 22, 23, 25, 26, and 27) would similarly challenge an 8 TF/s system, as would a suite of applications coming from ASCI[4] or from BBRP.
- **Programmatic requirements.** The programs and directorates invest in M&IC. Their ownership position allots to them a percentage of the machine based on size of investment. We recommend that the programs realize—for this procurement only—a bonus factor of 1.67 (the usual number is 1.35). For each dollar invested in hardware the programs would thus see $1.67 return in access rights, with the bonus taken from the institutional allocation. This gives the programs additional cycles to use at their discretion while still leaving substantial access to the institution. A highly desirable consequence of the allocation model is that institutional investments in computing are automatically steered to areas of high programmatic interest while stimulating the programs to invest more in M&IC. Any system procured needs to make room for this co-investment. Given recent investment levels, the programs will realize about 2.5 TF/s. Of course, many of these cycles will inevitably be allotted by the directorates and programs to the efforts described here, so care must be taken not to "double count" in sizing the system.
- **Extraordinary requirements.** We note that several of the projects (for example, DJEHUTY [Project ID 2], Fig. 2.5) have requirements that considerably exceed any system that we will be able to purchase in the near term, and several projects are ready today to utilize systems at the hundreds of TF/s level.

[4] As another example of extreme computing requirements that exist at this laboratory, the Opacity Project (not represented in our list of projects summarized here) stated at the ASCI PI Meeting held January 7–10, 2002, that it was ready for and could profitably use a petaflop/s system.

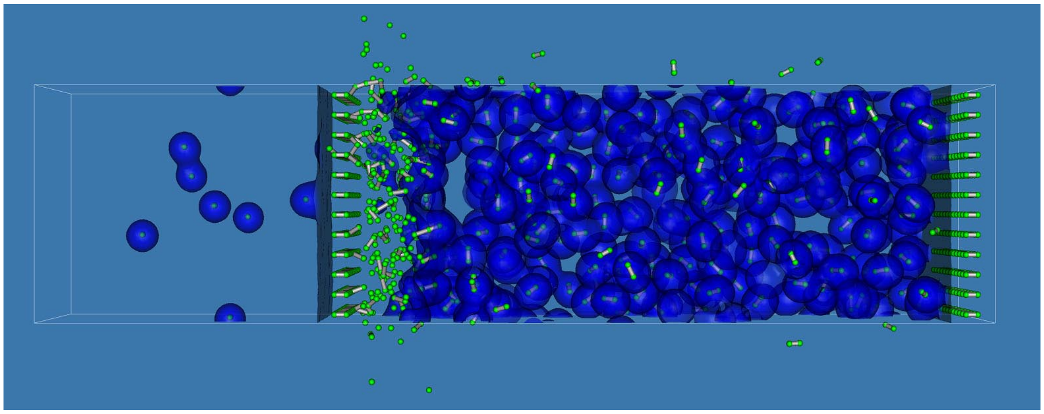

Figure 2.3. This simulation of the shock compression of deuterium is an important advance in understanding of material from first-principles molecular dynamics. This JEEP simulation shows macroscopic behavior from a microscopic model of matter and a strong correlation to observations of the actual phenomena.

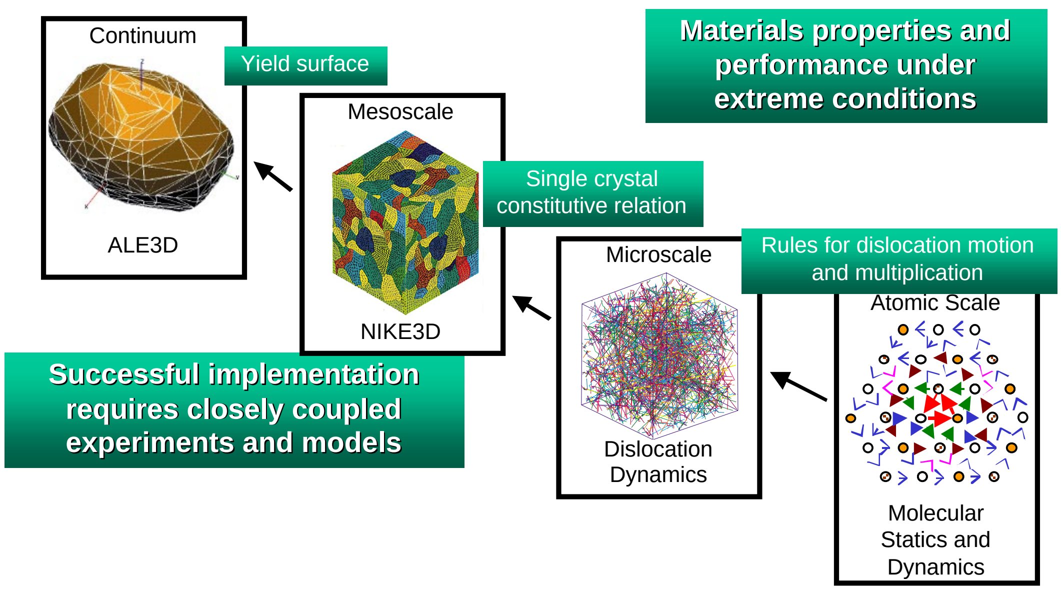

Figure 2.4. Multiscale modeling is needed to predict the properties and performance of materials. Materials science, computational biology, and the environment are obvious Grand Challenge candidates based on the user responses to the M&IC questionnaire.

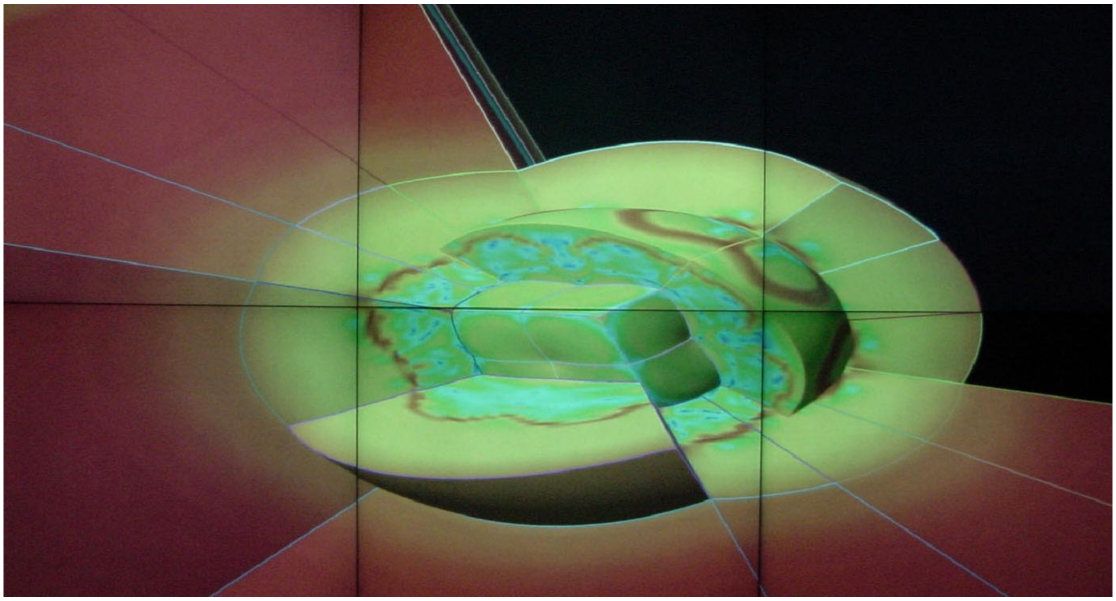

Figure 2.5. Improved determination of the size, age, and composition of the universe is a goal of the Djehuty project, the world's only operating code that can model a complete star in three dimensions. This picture shows a cutaway display of the simulation of a star on the Building 451 Power Wall.

3. Recommendations

In this paper, while we are focusing on the near-term requirements for immediate increases in capacity, we have kept two additional strategies in mind.

1. We are taking the view that the institution may be interested in at least one strategic computational thrust focused on a scientific goal that is of national interest. This could take the form of a Grand Challenge in computational physics, biology, or chemistry. We have satisfied ourselves, on the basis of the requirements stated by the 27 science team respondents, that such an effort requires at least 4 TF/s sustained per thrust *at this time* (scientists are requesting 8 TF/s). We are therefore looking at systems that will satisfy the needs of (1) a strategic thrust, (2) programmatic/directorate access through programmatic funding, and (3) institutional scientists working *outside* the scope of the strategic thrust. Taking into account overlaps between categories, this sums to about 10 TF/s.

	Grand Challenge	Programmatic Ownership	Other Institutional	Total
TF/s Required	4	2.5	4	10.5

2. We are also taking a longer-term view that LLNL would wish to take advantage of a unique opportunity to achieve distinction in unclassified computing by mid-decade. This could be achieved through a partnership with ASCI in 2004–2005 in which LLNL and ASCI work to bring a leading-edge, cell-based IBM system (BlueGene/L) to LLNL in the 200–400 TF/s regime. This would require an associated institutional applications effort beginning at the latest in mid-FY03 to target a strategic thrust to this system. This preparatory work would require a development platform, much like the computing systems that we propose in item 1 above.

These strategies suggest the following timeline:

FY02 and FY03—LLNL would procure a Linux system between 5.5 TF/s and 11 TF/s in FY02 and cover the costs primarily over FY02 and FY03 (cost estimates can be provided). This has the added bonus of putting LLNL in a position of rough equivalence with the most capable Office of Science and NSF sites. We emphasize that 5.5 TF/s is *not* an ample solution, and the resulting environment could soon be heavily contended, just as it is on ASCI Frost and TC2K today. If an 11-TF/s system proves too costly, an 8-TF/s system might still make all three classes of work as described in item 1 (above) possible.

FY04 and FY05—We recommend that the Laboratory consider a partnership with ASCI in which the institution invests to enhance the BlueGene/L system being pursued by ASCI. LLNL will receive a proportionate share of the ~200–400 TF/s system,[5] which is to arrive in FY04 or FY05. In the meantime, it is highly advisable that at least one LLNL science effort get started as a strategic effort (or possibly a Grand Challenge of national interest), in materials science, the biosciences, or the environment.

FY06 and FY07—It is likely that IBM will follow BlueGene/L with a petaflop/s-class system. LLNL should consider a partnership with ASCI to procure such capability for the institution.

[5] Whether the second processor in a cell can be used profitably for computation as well as communication determines whether this is a 200 or 400 TF/s system.

Appendix A—Fall 2001 M&IC Questionnaire

The following questions were submitted to major users of open Multiprogrammatic & Institutional Computing (M&IC) resources of the Livermore Computing Center (LC) in the fall of 2001. The individual responses are available as an addendum, "Unclassified Computing Capability: User Responses to a Multiprogrammatic and Institutional Questionnaire."

1. **Scientific intent and scope**
   Describe the project's scientific intention and its scope. Describe its importance to your program, directorate, or to the institution. What is the potential or current overall scientific impact of the proposed project?

2. **Project funding source, size in FTEs, PI, and major contributors**
   Describe the project's current funding source and rough number of FTEs on the team. Please name the PI or the main author and the major contributors.

3. **Code maturity level and history**
   Describe the current maturity level of the application code(s) used for this project, including level of parallelism, production status, where the code was developed, its history of use, what systems it has used in the past, and (if known) its level of efficiency at various node and processor counts. [The code should be checkpoint restartable—if the code does not currently have this capability, how long would it take to create this capability?] What capability does your project have to visualize the computational results?

4. **Anticipated scientific advance at 2–8 teraFLOP/s**
   What would be the scientific advance or breakthrough that could be achieved if this project received dedicated access to an M&IC Open Computing Facility capability platform of between 2–8 teraFLOP/s for 1–4 months? (Note: eventual size of the system is not yet determined, but would be at least 2 TF.)

5. **Anticipated scientific advance at 8–32 processor range**
   What would be the scientific advance or breakthrough that could be achieved if this project received copious access to banks of processors cooperating at the 8–32 processor range (25–100 GF)?

6. **Connection with experiment**
   How does this project support experimental science or how does this project depend on experimental data from previous (or planned) experiments?

7. **Requirements for scientific challenge at 2–8 teraFLOP/s**
   Assuming M&IC could provide your project a 2–8 teraFLOP/s resource for a specific scientific challenge, please describe your computational requirements for success, including (but not limited to):

   - Ratio of bytes of memory/flop (or how much memory per processor or how much total memory and how many total FLOP/s)
   - Ratio of bytes of disk/flop (or how much local and/or global disk would the run require)
   - What visualization requirements do you have? What visualization support would you require?

- Tertiary storage requirements in TB
- When you would be ready to start a 6–18 month project as described in the introduction?

8. **Requirements for scientific challenge at intermediate capacity level**
   Assuming M&IC could provide your project with "capacity at the intermediate capability level" (many 16–32 processor runs) for a specific scientific challenge, please describe your computational requirements for success, including (but not limited to):
   - Ratio of bytes of memory/flop (or how much memory per processor or how much total memory and how many total FLOP/s)
   - Ratio of bytes of disk/flop (or how much local and/or global disk would the run require)
   - What visualization requirements do you have? What visualization support would you require?
   - Tertiary storage requirements in TB
   - When you would be ready to start a project?
9. **Machine use pattern at 2–8 teraFLOP/s**
   Assuming the machine features 256–512 nodes with 2–4 processors with 3.6–4.0 gigaFLOP/s peak processors and 2–8 GB of memory and 80–160 GB of disk per node with a Quadrics interconnect (>300 MB/s), please describe how you would use this machine in the period in question in which you enjoy a preferential allocation (between 6 months and 18 months). Think in terms of the whole project effort as well as the capability science runs. What scaling or convergence studies would your project need to perform? Is there any experimental data that needs to be analyzed as part of the effort? What validation of the algorithms and/or application will need to be performed? This question requires some real (and realistic) thought, as we are using these requirements to size a system and make an honest case to the Deputy Director about the spectrum of mature unclassified efforts around the Laboratory.
10. **Machine use pattern at intermediate capacity level**
    Assuming the machine features many banks of nodes dedicated to 8–32 processor runs (think of the system architecture similar to that described in question 8 but dedicated to smaller problems, or think alternatively of TC2K dedicated to 8–32 processor runs), please describe how you would use such a system to achieve your goals. Think in terms of the whole project effort as well as the science runs. Is there any experimental data that needs to be analyzed as part of the effort? What validation of the algorithms and/or application will need to be performed? What is the approximate number of runs that need to be done and the approximate wall-clock time of each run?
11. **Memory vs compute trade-off**
    Would this machine be of use with less memory, say 1 GB per processor (.75 GB available to the user)? In other words, would you trade away memory to get more nodes or gigaFLOP/s?
12. **Comments on value of unclassified computing**
    Please take some time to describe the value of unclassified computing (including M&IC) to your efforts in the past and today. In particular, we would like to hear about programmatic impact. M&IC has not been an inexpensive investment, and we need to make the case to our

Deputy Directors that this investment has enriched and enhanced this institution. Questions we get concern “return on investment.” [For instance, has your involvement with M&IC allowed you to hire additional personnel? Has it enhanced your ability to meet programmatic milestones? Has it enriched the institution? (that is, have you been asked to give talks to review committees or in other forums in part based on your work with M&IC systems?)] Please dwell on unmet needs (especially in the arena of access and computer time) as well as needs that have been met adequately.

Appendix B—Response Summary

The following table provides a high-level summary of the responses received to the M&IC questionnaire. Twenty-seven responses were received that involve all nine of the technical directorates at the Laboratory.

The column headings used in the table are:

- Project ID—an arbitrary code phrase assigned to this project to allow for quick reference.
- Submitted By—the person or persons submitting the response. While the person submitting the response was often the projects principal investigator, this is not always the case.
- Title—the title of the project.
- FTEs—the number of full-time-equivalent employees directly involved in the project. Note that more individuals than the number of FTEs often work on a project as many people only contribute a fraction of their time to the particular project.
- Primary Funding—the primary sources of funding for the project.
- Supports Lab Programs—Laboratory programs supported by the project.
- Team From—Laboratory organizations of the primary project team.
- Also Involves LLNL Orgs—other Laboratory organizations that are closely involved in the project.
- Involves External Orgs—organizations external to the Laboratory that are directly or closely involved in the project.

	Project ID	Submitted By	Title	FTEs	Primary Funding	Supports Lab Programs	Team From	Also Involves LLNL Orgs	Involves External Orgs
1	ALPS	Milo Dorr	ALPS (Adaptive Laser Plasma Simulator) Project	3.15	LDRD, ASCI	ASCI, NIF, LDRD	Comp/CASC	Comp/CAO, DNT/X, PAT/M	
2	DJEHUTY	Dave Dearborn, Don Dossa	The Djehuty Project	2			Comp/CASC, PAT/V		
3	AMRh	Allen Kuhl, Jeff Greenough	AMRh Code Development Project	4.5	ASCI	A Program	DNT/A	DNT/X	
4	Fermion MC	Malvin Kalos	Physics Problems in Very High Dimensions	2	ASCI, LDRD	ASCI, A Program, LDRD	DNT/A, PAT/H, Comp/CASC		U. Trento
5	DD-ICF	Steven Langer	Direct-Drive ICF	1		ICF	DNT/X		U. Rochester
6	Z3	Barbara Lasinski, Bruce Langdon, Bert Still	Z3 Project	1	LDRD, NIF	LDRD, NIF, A Prog	DNT/X		DOE Fusion Energy, collaborations with universities and other labs
7	Mat-Shock	Brian Wirth, James Stolken, Maria Caturla	Shock Propagation in Materials	2.5	ASCI, LDRD, NIF	ASCI, NIF, LDRD, MFE	C&MS/MSTD	DNT, C&MS, Eng, PAT, NIF	Basic Energy Science
8	Mat-Rad	Tomas Diaz de la Rubia, Brian Wirth, Bill Wolfer, Maria Caturla	Radiation Damage of Materials	3.5	DOE-DP, DOE-FE, ASCI	ASCI, other	C&MS/MSTD	CMS, PAT/FE, Eng	DOE Fusion Energy and Nuclear Energy Plant Optimization (NEPO)
9	Cell Modeling	Carl Melius	Modeling of Biological Cells for Bioterrorism and Health	3	LDRD	LDRD	C&MS		DOE Office of Science - ASCR
10	Biochem	Andrew Quong, Chris Mundy	First-Principles Molecular Dynamics for Understanding Fundamental Biochemical Interactions	3	LDRD	LDRD	C&MS/ANCD		
11	CompBio	Mike Colvin, Felice Lightstone	Computational Biology Project	4	LDRD, DOE-SC/OBER, DOE-SC/SciDAC	BBRP, LDRD, other	BBRP	Comp/CASC, PAT	
12	GFMD	Lin Yang	GFMD (Greens Function Molecular Dynamics) Project	3	ASCI, LDRD	ASCI, LDRD	PAT/H	PAT, C&MS	Wright-Patterson Air Force Base
13	BOUT	Xueqiao Xu	BOUT (Boundary-plasma Turbulence) Project	2	MFE	MFE	PAT/FE		
14	NuclStruct	Erich Ormand	Ab Initio Nuclear Structure	2	LDRD, ASCI	PDRP, LDRD, ASCI	PAT/N		U. Arizona, Iowa State, U. Washington, San Diego State, CNRS – Strasbourg
15	JEEP	Giulia Galli, Francois Gygi	The Jeep and Quantum Simulations Projects	16	LDRD	LDRD, ASCI, BBRP, other	PAT/N, Comp/CASC	BBRP, NAI	several international collaborations with universities and other labs
16	PHENIX/HBT	Ron Soltz	PHENIX/HBT	11	LDRD, DOE-SC	LDRD, other	PAT/N		Brookhaven NL
17	MD3D	Jim Belak, Robert Rudd	Microscopic Origins of Dynamic Fracture Project (MD3D)	3	ASCI, LDRD	ASCI, LDRD	PAT/N	PAT, C&MS, Eng	CNRS-Grenoble
18	pF3d	Bert Still, Ed Williams, Richard Berger, Bruce Langdon, Laurent Divol	pF3d - Predictive Laser-Plasma Interaction Modeling	4	NIF	NIF	DNT/X, PAT/M	NIF	U. Rochester (LLE), Princeton
19	NIF gas	Steven Sutton	Gas Distortion Characterization for the National Ignition Facility	2.8	ASCI	NIF, ASCI	Eng/LSED		
20	EIGER codes	Rob Sharpe	Electromagnetic Effects in High-Frequency Integrated Circuits	3	LDRD, Eng, WFO (DARPA, DoD)	LDRD, Eng	Eng/DSED	Comp/CASC	Several universities, UC Davis, DARPA, DOD HPCMO
21	NDE	Pat Roberson	Nondestructive Characterization	30	DOE-DP/ESC, NIF, other DP, WFO (FAA, Knolls Atomic Power Laboratory, others)	Eng, ESC, NIF, others (FAA, KAPL)	Eng/DSED		
22	E3D	Shawn Larsen	Seismic and Acoustic Wave Propagation	10	NAI (NNSA-NN), DOE-FE, LDRD, WFO, NIF	EED, NAI, Eng, NIF, LDRD, others	Comp/CAR, EED/GGS, Eng/CCDS/DSED/EETD,	NAI, PAT, NIF	University collaborations, USGS, NSF, IGPP, DOD, licensing agreement with petroleum industry, USBR (b. recl)
23	HP-CFD	Bob Lee, Don Ermak, Stevens Chan	High Performance Computational Fluid Dynamics (CFD) Models	3	NAI (NNSA-NN), LDRD, DOE, SANDIA	EED, NAI, LDRD	EED/ASD	NAI	DOD, DOE, Sandia, Arizona State U., Los Alamos
24	NUFT-C	Bill Glassley, John Nitao	NUFT-C Project	10	LDRD, DOE (?DOE-SC?), WFO	LDRD, EED, other	EED/GET		collaborations with petroleum industry
25	HR-GCS	Phil Duffy	High-Resolution Global Climate Simulations	1	DOE-SC/OBER	EED, other	EED/ASD	Comp/CASC	
26	AtmosChem	Doug Rotman, Cyndi Atherton, Peter Connell, Cathy Chuang, Jane Dignon, Dan Bergmann, John Tannahill, Philip Cameron-Smith	Global Atmospheric Chemistry	12	DOE-SC, NASA	EED, other	EED/ASD		DOE-SC, NASA, NCAR, LANL, ANL, ORNL, PNNL, LBNL
27	Earthquake	Charles Noble	Morrow Point Dam Earthquake Analysis	1	Eng. WFO	Eng. Other	Eng/NTED		U.S. Bureau of Reclamation

Appendix C—Project Needs

The following table provides a high-level summary of the computational needs of the projects that responded to the M&IC questionnaire. The needs summary is divided into *capability requirements* (needing the bulk of a computing resource of ~8 TF) and *large-capacity* requirements (needing time on a computing resource of ~100 GF).

The column headings used in the table are:

- Project ID—an arbitrary code phrase assigned to this project to allow for quick reference.

Separately, needs for Capability Requirements (8 TF/s) and Large-Capacity Requirements (100 GF/s) are listed for:

- time needed—computed time needed in terms of number of runs, duration, and/or size of resource.
- mem/node—amount of main memory needed per compute node.
- disk—amount of scratch disk storage needed during active runs.
- archive—amount of archival storage needed to store results.
- when ready?—when project will be ready to utilize this resource.

	Project ID	Capability Requirements (8 TF/s) time needed	mem/node	disk	archive	when ready?	Large-Capacity Requirements (100 GF/s) time needed	mem/node	disk	archive	when ready?
1	ALPS	several 2-day runs	8 GB per node **would be tight!** but probably sufficient.	8 TB per run (16 GB local disk per node)	8 TB per run	Apr-02	100s of 12-hour runs	>~1 GB per processor (like TC2K)			NOW
2	DJEHUTY	>~30 1-wk runs		3.5 TB /run							
3	AMRh	multiple runs, 4 wk total	>~1 GB per processor	100's GB per run	several TB	NOW	multiple refinement and debugging studies	>~1 GB per processor	10's GB per run	several 100 GB	NOW
4	Fermion MC	few months +++	100 MB per processor	0.5 TB per run (1GB per node)	a few TB	in 30 days	10 50-hr runs per study	100MB per processor	16 GB per run (1GB per node)	a few TB	in 90 days
5	DD-ICF	(no time requested, now - don't anticipate need for 1-2 yr)				in 1-2 years	many 32P calculations	HYDRA ~0.5 GB per processor, CRETIN on ES45	60-70 GB per run, 60 runs/yr	4 TB for 60 runs per year	NOW
6	Z3	17-20 days per run, multiple runs (10 runs over 2 yr)	>~1 GB per processor	50 GB per concurrent run	500 GB per run	in 180 days ("a few months")	3-7 days per run, many capacity runs per 2-8 TF capability run	>~1 GB per processor			in 90 days
7	Mat-Shock	spherical shocks - 144 h @256 P, structural transforms - 384 h @256 P, shock and defects - 240 h @256 P	<~0.5 GB per processor	>30 GB	2-3 TB	NOW					
8	Mat-Rad	85 keV recoils - several runs @200 h @256 P, damage overlap - 60 runs @20 h @256 P, interaction mechanisms - 10 runs @350 h @1024 P	<~0.5 GB per processor	>30 GB	2-3 TB	NOW					
9	Cell Modeling	150 hr/yr of whole machine, 20-30 hr/yr of partial machine - starting in FY03				FY03					
10	FP-Biochem	250 hr/yr of total machine, 30-40 hr/yr of partial machine - starting in FY03				FY03					
11	CompBio	4 months for 1 project - ready NOW - plenty of systems to run, limited only by CPU time	min 1GB /node - 2 GB /node better	1 TB gpfs needed, >>1 TB desired	5 TB	NOW	10-15 one-month runs per system to be modeled, multiple systems (10 systems over next year) can be modeled - ready NOW	1 GB /node	2-5 GB /run		NOW
12	GFMD	initial studies - 40 different configurations @2 wk /config, milestone calculation will run 2 mo	2-3 GB /node	1000* 5 GB => 5 TB	5 TB	Jul-02	24-48 hr/run, need 1000 runs	4-8 GB /node	25-100 GB /run	10 TB	NOW
13	BOUT	100 runs, ~10 hr /run	1.5 GB per processor	<~1 TB	<~ 100 TB	early CY02	100 runs, ~10hr/run				
14	NuclStruct		2 GB per processor	0.5-2 TB		NOW	many 1node runs	1 GB/P, 4P /node	1 TB		NOW
15	JEEP		>~ 2 GB /node (0.5-1 GB/GF)	1 TB gpfs + 20 GB /node local	3-5 TB	NOW		>~ 2 GB /node (0.5-1 GB/GF)	1 TB gpfs + 20 GB /node local	3-5 TB	NOW
16	PHENIX/HBT						PISA - 5-10 @1wk, SIC - 12 wk @25 GF (8P), CRAFT - 12 wk @25 GF (8P)	4-8 GB /node	10-20 GB	2 TB	NOW
17	MD3D		small <1 GB	1-10 TB		Spring '02	**4-mo @32P (7x1-wk runs, 50x1-da runs)**	small <1 GB	1-10 TB		
18	pF3d	1 mo /calc -- 3 major calcs (over next year)	1000 GB total => ~2 GB /CPU	1-2 TB local (restarts), 100 GB global (history)	10's*1 TB	NOW	5h for 165 TF (10% eff), 100 times	22 GB total => >~3GB /node	0.5 TB	5 TB	NOW
19	NIF gas	not available, scaling and timing studies needed				? - assume ready in 6 months	not available, scaling and timing studies needed				? - assume ready in 6 months
20	EIGER	several 3-week runs	>1 GB per node	~100 GB		~10/02		>1 GB per node	~100 GB		~10/02

	Project ID	Capability Requirements (8 TF/s) time needed	mem/node	disk	archive	when ready?	Large-Capacity Requirements (100 GF/s) time needed	mem/node	disk	archive	when ready?
21	NDE	1-5 days/run, 20 runs/yr	>~1 GB per node, now; ~2 GB per node, later	~0.5-2 TB per run	n*3 TB	? - assume ready in 6 months					
22	E3D	12-48 h/run, 10-40 s of runs - ideal for dedicated time at several intervals, say, six 1/2-month allocations over 12 months	2-4 GB /node	0.2-1 TB	<~1 TB	NOW	4-8h/run, 500-1000 runs over 3 mo - ideal for dedicated time at several intervals, say, two 2-month allocations over 12 months	1-2 GB /node	0.1 TB global disk	<~1 TB	NOW
23	HP-CFD	1/2 machine every couple of months, 2 wk/run	>1 GB per node			May-02	2 runs/wk, 2days/run	>1 GB per node			NOW
24	NUFT-C	100 h/run, multiple runs	2 GB /node	0.5 TB	10 TB	Jan-02	20h/run, multiple runs	1-2 GB /node	0.1TB	1 TB	NOW
25	HR-GCS	several simulations per year at 3*10^5 P-hr	"like Frost" => 8 GB /node?	0.1 TB local disk	5 TB	NOW	twenty calculations per year at 5*10^4 P-hr	guess several GB /node needed?	0.1TB	<1 TB	NOW
26	AtmosChem	12 studies @1/8-1/4 machine @2 wk, whole machine 1-4 mo	0.5B/F => 16 GB /node	5-10 TB	5 TB	NOW, for some studies - largest simulations ready ~6/02	3-6 wk/run, multiple studies	4 GB /4P node			NOW
27	Earthquake	3 wk/Q	8 GB per node	100 GB	200 GB	NOW	3 wk/Q	8 GB per node	100 GB	200 GB	NOW

Appendix D—Scientific Advances

The following table provides a high-level summary of the scientific advances that these projects anticipate they could achieve with access to additional M&IC computational resources.

The column headings used in the table are:

- Project ID—arbitrary code phrase assigned to this project to allow for quick reference.
- Using Capability Platform (8 TF/s)—scientific advances that can be achieved using a new *capability platform*.
- Using Large-Capacity Platform (100 GF/s)—scientific advances that can be achieved using a *large-capacity platform*.

Multiple advances are envisioned for many of the projects. Boxes that are blank and shown shaded in gray indicate projects that anticipate no need for that particular resource.

	Project ID	Using Capability Platform (8 TF/s)	Using Large-Capacity Platform (100 GF/s)
1	ALPS	Full-scale 3D runs at the size of the experimental configuration currently being used on the Omega facility at the University of Rochester in support of the NIF program.	This would allow us to establish a qualitative understanding of the scaling of beam energy transfer with respect to a large number of parameters with a large number of scaled-down 2D problems covering a wide range of plasma conditions and beam configurat
2	DJEHUTY	Study rapidly rotating and binary stars, stars with short lifetimes (MACHO data); resolve 3He overproduction anomaly.	
3	AMRh	First fully resolved compressible shock-driven flows in 2D; 3D turbulent breakup of vortex rings (Omega); AWE supersonic jet experiment.	Lower resolution runs of the 2D studies under 2-8 TF platform (fill out the scaling studies), investigate equation-of-state (EOS) issues, and various sensitivity issues.
4	Fermion MC	Simulate the water dimer by an exact all-electron method.	Ability to do clusters of water molecules, plus hundreds of atoms of He-3. The physical chemical properties of water are at the heart of much chemistry and biology.
5	DD-ICF	Model the full surface of a direct-drive capsule, resolving all the relevant perturbation wavelengths. Such a calculation would be undertaken only after issues under 0.1T capacity calculations are resolved.	Answer various questions about direct drive implosions: 1) Does the capsule shell break up during the implosion? 2) Which wavelengths contribute most strongly to the RMS perturbation of the interface between the fuel and the surrounding shell? 3) How should the ALE package in HYDRA be tuned to best resolve the imploding capsule? These questions are best addressed by running simulations covering only portions of the capsule surface.
6	Z3	Modeling a complete f/8 speckle, the relevant plasma volume of a NIF beam, for an interesting period of time, ~ 100ps, in terms of laser-plasma interactions.	Many large-scale 2D simulations. Simulations in 2D of a short-pulse high intensity laser propagating in a plasma which spans electron plasma densities from underdense to overdense would allow us to begin to assess how beam focusing and breakup in the underdense plasma affect the subsequent production of directed particle beams. To date, very little has been done on this key question.
		Scale up to larger volumes and perhaps more than one f/8 speckle. At speckle size one and greater, we could then for the first time model beam propagation with a first principles code for relevant space and time scales of a realistic plasma.	We would investigate algorithm enhancements such as collision and ionization models as well as more elaborate boundary conditions. These are necessary studies as we seek to expand the scope of traditional PIC modeling.
		Scale our 2D simulations in the overdense region to 3D and provide a first principles prediction on the directed angular distribution and energy spectra of the charged particles produced in these laser matter interactions. Production and transport of these particle beams is a key component of the fast-ignitor scenario and other applications.	
		For the NIF parameter regime, experiments now can isolate a single speckle and unprecedented direct comparisons on beam propagation and instability saturation could then be made with the PIC first-principles description. In the history of PIC modeling, experiment planning and analysis have always been essential components. However, the smaller space and shorter time scales of the simulations have historically meant that there needed to be one or more intermediate steps in relating the PIC results to experiment. We will be able to reduce or remove these intermediate steps.	
7	Mat-Shock	Spherical shocks in the presence of grain boundaries (nanocrystalline materials): we will study the effect of a grain boundary in the propagation of a shock by including small grains in the material (~20 nm). We will look at the effect of grain boundary orientation (low and high misorientation between grains), shock pressures (at least 4 different values) and materials (at least 2 materials).	
		Structural transformations induced by laser-driven shocks in different materials and the effect of surfaces: large system sizes will be necessary for these simulations to compare directly with experimental studies. Two materials will be tested and 4 shock pressures and 4 pulse shapes.	
		Shock interaction with defects produced during self-irradiation: in this case different shock pressures and defect types will be tested. We estimate at least 48 cases will be necessary (4 pressures, 3 defect types, 4 sizes).	
8	Mat-Rad	Identify primary damage in Pu due to 85keV recoils (self-decay): an α-decay in Pu produces a recoil of 85keV in energy. We will be able to study the damage produced by such recoil with interatomic potentials developed specifically for Pu.	
		Determine damage overlap and high dose defect production in glasses & metals	
		Determine defect-grain boundary interaction mechanisms (important to understand denuded defect zones observed in metals and the potential radiation resistance of nanocrystals)	
		Quantify the impact of defect clusters on mechanical strength over a wide range of defect cluster types and materials systems	
9	Cell Modeling	Test the 3D reaction-diffusion capability of ALE3D by modeling human epithelial cell tissue. We will investigate the transport of simple molecules across the cell and its membranes, along with the signaling proteins within the apical and basolateral membranes regulating these transport processes.	
		Use the object-oriented, modular framework to test the complex chemical species built up from the protein machines. We will use the chemotaxis process of microbial cells as our test case for treating these complex chemical objects.	
10	FP-Biochem	Understanding ligand-receptor interactions of Tiron and beryllium in support of experimental efforts. In order to obtain an accurate picture of the energetics of receptor-ligand binding as it occurs in the body, the inclusion of solvating water is imperative. This will take us into the state-of-the-art in terms of number of atoms simulated by first-principles molecular dynamics.	

	Project ID	Using Capability Platform (8 TF/s)	Using Large-Capacity Platform (100 GF/s)
		Studying prototype molecular targeting agents that are studied in the laboratory. The large number of atoms that are needed to describe these systems render standard methods of first-principles molecular dynamics useless. In order to circumvent this problem we have proposed a methodology that will allow us to use both accurate electronic structure in the region where bonding is important (e.g. the metal-ligand interaction) and approximate quantum mechanical models to treat the large number of atoms in the protein scaffolding.	
		Study of a crystal of dimerized DNA bases via a first-principles approach to help elucidate experimental data on the effects of the condensed phase on radical stabilization. This isis aimed at elucidating a molecular picture of radiation damage to DNA bases.	
11	CompBio	The model of an Ape1 active site could be simulated, which would provide important information about the mechanism of phosphate hydrolysis in a realistic environment. We would be create a model enzyme active site and include a small piece of DNA to be cleaved.This simulation would be among the largest first principles molecular dynamics calculation performed to date. 500-600 atoms	We could efficiently perform many simulations of systems in the range of 200-300 atoms in size. Computer resources of this type would be ideal for a number of the projects we are currently working on. For example, in the calculation of a free energy profile for a reaction in solution we need to perform 10 to 15 independent simulations along a reaction coordinate. Each of these simulations need to be 3 to 4 ps in length and could be computed with approximately one month of dedicated access per simulation (all simulations can be run simultaneously). For a project such as the activation of phosphoramide mustard anti-cancer drugs this would enable us to develop a deeper understanding of how variations of the drug affect its activity and could be directly verified by experimental measurement.
12	GFMD	At 8 Tflop/s, GFMD will be able to simulate materials with 10 million atoms (a 10-fold increase from our current maximum), allowing for direct comparison of simulation material-strength results with experiment. After the code is successfully validated through comparisons with existing experiment, GFMD will then be used for predictive simulations for extreme conditions of temperature and pressure that are not now accessible to experiment .	We will run a large number of simulations for various dislocation configurations such as kink formations. This study would allow us to build accurate data-base for screw dislocation mobility model that is the key input into much larger-length-scale simulations. Using these resources would allow us to have 2 data points per day to meet the future milestones set for the LLNL program on multiscale modeling of strength and failure of materials.
		Simulate a full-scale 3D modeling of dislocation junctions that have been observed in TEM experiment. This type of capacity will move our simulation capability as the world-class leader in the field of atomistic simulations of dislocation-dislocation interactions and multi-scale modeling of strength of materials. This type of simulations is an important step to have direct comparison to experiments.	
		3D calculations of dislocation-dislocation interactions that involve 2-3 millions atoms distributed over 500,000 structured cells in order ot get meaningful physics out of our simulations. We have been running some scaled-down (~200,000 atoms) 3D calculations on TC2K and IBM SP P3 machines to calibrate the scalability and accuracy of our calculations.	
		Simulate a full-scale 3D modeling of dislocation junctions that have been observed in TEM experiment. This type of capacity will move our simulation capability as the world-class leader in the field of atomistic simulations of dislocation-dislocation interactions and multi-scale modeling of strength of materials.	
13	BOUT	Incorporation of neutral transport and the associated neutral effects on both plasma turbulence and the plasma particle, momentum, and energy balances in BOUT will allow a critical investigation of the tokamak "density limit".	Investiagte the L-H ("low" to "high") transition for a simpler magnetic geometry, such as circular cross section tokamaks and scaling calculations. The evolution of the plasma profiles during the time of a L-H transition can be followed within a turbulence code that evolves the background profiles.
14	NuclStruct	Our primary focus would be to carry out production runs for the suite of nuclei ranging from mass eight to sixteen. Our codes have already been validated on smaller machines using smaller basis spaces. Scaling tests would have to be performed on the massively parallel machine.	Would use as an important to aid with 2-8 TF runs.
		Dedicated access to 2-8 TF would enable us to extend the our first-principles nuclear structure calculations across six major shells, which would improve convergence from 500 keV to approximately 200-300 keV and guarantee our ability to complete calculations for all relevant nuclei from boron-10 up to oxygen-16.	
15	JEEP	After initial scaling runs, the machine would be filled with one large problem (128-256 nodes) and 3-4 smaller problems (32-64 nodes).	Several 32-node jobs would be run simultaneously. This is an efficient use of resources in some problems, notably in the calculation of the free-energy profiles of some biochemical reactions, which involve several independent molecular dynamics runs. The duration of the runs is typically limited by the batch system quota. If the machine (or part of it) is dedicated, job duration is limited by the machine's MTBF. (On Frost, several JEEP runs of >10 hours were completed). The signal-catching capabilities of JEEP allow for efficient use of multiple chained jobs with minimal user intervention.
		High Pressure Fluids (HPF)--Microscopic simulations of shock propagation in high explosive mixtures (at 2-8 TF). These sub-projects could have huge programmatic impact for DNT (both A and B division) and for NIF.	High Pressure Fluids (HPF) -- Microscopic simulations of shock propagation in heavy metals (at 25-100 GF). These sub-projects could have huge programmatic impact for DNT (both A and B division) and for NIF.
		Semiconductor Nanostructures (SN)--Optical properties of dots in solution on the fly (at 2-8 TF). These sub-projects could have a huge scientific visibility and eventually will lead to close collaborations with NAI.	Semiconductor Nanostructures (SN)--Virtual (real time) atomic manipulation of semiconductor nanostructures in solution (at 25-100 GF). These sub-projects could have a huge scientific visibility and eventually will lead to close collaborations with NAI.
		Computational Biology (CB)--DNA base pairs in solution (at 2-8 TF). This sub-project would have as well a huge scientific visibility.	Computational Biology (CB)--Drug attack to DNA in solution (at 25-100 GF). This sub-project would have as well a huge scientific visibility.
16	PHENIX/HBT		**PISA code needs**--Generating simulation data with PISA occurs in batches of 100k events, each requiring 100 GF (32 processors) for a period of 1 week. We estimate needing to perform approximately 5-10 such runs per year.
			Source Imaging Code needs--including initial systematic studies, and a series of ten data sets, we would need access to 25 GF (8 processors) for a total integrated time of 12 weeks.
			CRAFT code needs--At this time, we estimate computing requirements for CRAFT to be comparable to those of the source imaging code.

	Project ID	Using Capability Platform (8 TF/s)	Using Large-Capacity Platform (100 GF/s)
17	MD3D	**Growth of a 10 nm void in Tantalum at strain rates of 10^8-10^9/s.** BBC metal Tantalum is a major focus of the Dynamics of Metals program. The appropriate interatomic force models for Ta are 10-100 times more computationally intensive than for simpler metals such as Cu.	**Map out the dislocation mechanisms of void growth as a function of triaxiality.** Preliminary studies show that the dislocation mechanisms of void growth are sensitive to the stress state, i.e. relative amount of shear stress to isotropic tension, this is known as triaxiality.
		Simulate the behavior of a 100 nm void under tension (about 1 billion atoms) at rates of 10^8-10^9/s. We expect significant size effects and need to quantify the dislocation processes for voids larger than 10nm we can do now. This represents a 10 times increase in volume; 1000 times increase in number of atoms.	**Map out the dislocation nucleation as a function of void roughness.** The rough regions concentrate the applied stress. We have assumed the initial voids to be perfectly spherical.
		Simulate the behavior of a 100 nm void under tension (about 1 billion atoms) at rates of 10^8-10^9/s.	
18	pF3d	**Pf3d is our primary tool to bridge between (largely theoretical) studies of microphysics and macroscale LPI experiments. However, to connect the 2D calculations back to the experimental results requires 3D "capability" calculations.**	As we improve our microphysical models--for instance how do the ion acoustic waves involved in SBS saturate--we incorporate distillations of them in pF3d and then see of what range of parameter space they can model experimentally. This requires relative large numbers of "capacity" calculations (in 2D) to make progress.
		With access to 2-8 Tflop/s for 1-4 months, we could perform the first simulation on the entire Nova beam volume, and perhaps even the volume of a NIF outer beam.	Electron thermal conduction must be better understood. By its nature, electron heat conduction in a laser driven hohlraum is nonlocal and nonlinear, making a computational model difficult. This can be studied in 2D using the 25-100 GF/s system.
		We would compare the results of these full beam simulations to the smaller corresponding letterbox calculations and determine just how well, and when, we can extrapolate. These "capability" calculations on 2-8 Tflop/s system would allow us to make more meaningful use of "capacity" resources.	Saturation mechanisms for both stimulated Brillouin backscatter (SBS) and stimulated Raman backscatter (SRS) can be studied in 2D using the 25-100 GF/s system. SBS is generated by the interaction of the laser pump with an ion acoustic wave, while SRS is generated by the interaction with an electron plasma (Langmuir) wave. Part of understanding these saturation mechanisms involves parameter studies with pF3d, in which we vary the frequency of the light, the power in the beam, beam models (with and without smoothing techniques), and plasma conditions. We can directly compare pF3d simulation results with results from the particle-in-cell code ZOHAR (fully kinetic, relativistic electromagnetic in 2½d)--or its relative Bzohar (mobile ions, but Boltzmann fluid electrons), or its daughter code Z3 (3D MPP fully kinetic code)--for small problems (e.g., a single speckle) to validate the saturation models. These comparisons will require many "small" runs (25-200 Gflop/s) in 1D and 2D and some in 3D (with PIC codes and pF3d) to do parameter studies in a single speckle. Then larger (2-8 Tflop/s) pF3d simulations would be used to study the impact of those models on longer plasmas with RPP beams (when many speckles are present).
19	NIF gas	Mesh sensitivity studies, which have been restricted to date because of the machines available. It is imperative that these be performed prior to analyzing the larger structures on NIF.	This type of capability would allow us to run simulations for smaller geometry portions of the NIF system in the execution clock-times required. For larger geometry regions, this would still be of use but would reduce the impact that analysis would have on commissioning decisions.
		Development of a distortion database for systematic thermal conditions. The information in this database will be used in NIF propagation simulations.	
		Analysis of non-systematic thermal disturbances that are identified during commissioning. This would include the effects of spot heat sources in the transport and switchyard beamtube areas.	
		The shot-rate of NIF is going go be largely dictated by optical distortions from thermally driven convection currents. In addition, additional distributed heat sources in the system could lead to significantly increased focal spot size. It is important for project success that a computational capability be developed and exercised to allow simulations of convection currents to be performed in a day (they currently take several weeks). This type of capability would allow us to perform calculations during NIF commissioning to address issues as they occur.	
20	EIGER	It is extremely difficult, if not impossible, to measure the electromagnetic fields within an IC, this project would have a tremendous impact on our understanding of electromagnetic effects inside of high frequency ICs.	Analysis of electromagnetic coupling of adjacent microstrip interconnects could be studied in great detail
		Will enable extremely detailed simulation of entire ICs	Electromagnetic coupling of adjacent inductors in a microwave low-noise amplifier
			This type of research is extremely interesting to the electronics designers in the DOE, DOD, and industry.
21	NDE	Ability to nondestructively characterize objects and materials at resolution and detail that has not previously been possible.	
22	E3D	A wide range of diverse scientific problems that can be addressed if such resources are available. Capability computing at the 2-8 TFlop level will satisfy needs in: 1) underground structure detection, 2) nuclear nonproliferation; 3) oil exploration; and 4) earthquake hazard analysis.	A wide range of diverse scientific problems that can be addressed if such resources are available.
		In the case of underground structure detection and nuclear nonproliferation, it is desired to increase the source frequency by a factor of between two to four. In the case of oil exploration and earthquake hazard analysis, it is desired to decrease the geologic velocity by a factor of between two and four. Such simulations will require machine capabilities that are at least 50 to 100 times the 10-20 GFlop simulations that are preformed today (i.e., 0.5-2 Tflop). In general, seismic wave propagation problems are computationally limited to lower frequencies and higher geologic velocities (e.g., rock instead of soft sediment). Doubling the source frequency or reducing the geologic velocity by a factor of two requires an eight-fold increase in memory and a sixteen-fold increase in compute time.	Future seismic modeling problems, particularly those involving underground structure detection and oil exploration, likely will require hundreds if not thousands of simulations where the geologic model and other simulation parameters are slightly perturbed. This is necessary to accurately characterize the subsurface in the presence of fine scale geologic heterogeneity. For example, the detection of underground structures is an extraordinarily difficult problem. It is possible to solve this problem with a preponderance of data or a preponderance of simulation. It is likely that very limited data will be available in real world applications, hence the need for vast intermediate capability computer resources. Each relatively small 3D simulation needed for this type of analysis currently requires several hours of compute time using a few dozen processors. It is anticipated that 500 to 1000 simulations could be performed over a three-month period with dedicated use of 32 processors.

	Project ID	Using Capability Platform (8 TF/s)	Using Large-Capacity Platform (100 GF/s)
23	HP-CFD	We would use the machine to make building scale flow and dispersion calculations with very high resolution (say, ~1 m grid spacing) and advanced turbulence model such as LES (large eddy simulation) to model atmospheric turbulence more accurately. Such results could be used to demonstrate the Lab's unique capability in this field and attract more funding from potential sponsors.	We have been using similar computing capability for some time and expect to require continued support from M&IC in the future.
		We would be in a position to solicit funding via performing a number of breakthough calculations of the 10 to 100 million grid point range which would be considered state-of-the-art.	It would enable us to perform numerous parameter studies and further improve our models in both numerics and physics.
24	NUFT-C	There are diverse suites of problems that could be addressed by this code, each with its own unique set of requirements. For most of these problems, we are run-time limited; it is common to have the need to simulate thousands of years of processes, but be time-step constrained by the relatively short time constants affecting transport processes.	Such activities would represent a scaled down version of the efforts described for 2-8 Tflop/s platform. Focus wold be on examination of specific chemical interactions and impacts.
		High level nuclear waste disposal relies on the ability to simulate the thermal-hydrological-geochemical evolution of the potential repository site for thousands of years into the future. Currently, such simulations are simplified (limited number of chemical species considered at limited resolution) because of the heavy computational burden associated with such calculations. Consequently, there has never been a full simulation of a potential repository in which a simulated release of radionuclides from a leaking waste container has been represented.	To date, there have been only a handful of NUFT-C runs that consider the repository behavior (thermal-hydrological-geochemical effects in the natural system without radionuclides). Many scenarios have yet to be evaluated in which different heat loadings and tunnel geometries are considered for their impacts on performance. Copious access to banks of processors would allow a dramatic and systematic improvement in the ability to optimize a repository for overall safety and performance.
		Conducting such a calculation would be a breakthrough in the development of safety arguments for nuclear waste repositories. It would allow clear evaluation of the uncertainties associated with previous simplified and abstracted models, would allow unambiguous representation of the interplay between coupled processes that influence contaminant transport, and provide a rigorous description of the rates and three-dimensional flow pathways released radionuclides would follow.	
		This accomplishment would be an important breakthrough by addressing an important national issue (radioactive waste disposal) that impacts nuclear materials management. The currently most advanced simulation tool for this purpose (NUFT-C) has focussed on simulating the evolution of the natural system, perturbed by the thermal impact of emplacing nuclear waste. These simulations have taken upwards of 80 hours on an IBM-SP 2, utilizing up to 256 processors. Consideration of radionuclide transport, which was not treated in these simulations, would increase the problem size by approximately two orders of magnitude.	
		The primary challenge in these simulations is constructing a conceptual model that realistically represents the system being considered, but which also honors the limited data sets available that describe physical parameters that apply to the complex natural system being considered. Generally, these simulations are data starved. The general strategy is to construct a model based on the conceptualization and run numerous preliminary simulations to refine those initial and boundary condition estimates that had to be made in the absence of hard data. This is often the most time consuming aspect of running these simulations, and can take weeks. Once simulation initial and boundary conditions are satisfied, the simulations are run and results compared to available data.	
25	HR-GCS	The answer to this depends on how well our code scales on this machine. If, as expected, our code scales poorly, we could efficiently use large numbers of processors by running an ensemble of high-resolution calculations (ed: i.e., multiple capacity calculations?). Because of the chaotic nature of the atmosphere, an ensemble of calculations is needed to obtain the most accurate predictions. (We could, for example, run an ensemble of seasonal forecasts predicting the effects of the upcoming El Nino.) Of course, if our code did scale well on the new machine (which I don't expect) then we could push to even higher spatial resolutions. The preliminary high-resolution climate simulations we have already performed mean that we are ready to do large-scale production runs now. (Indeed, we are already doing such runs on Frost and other machines.) No evaluation of algorithms, etc. would be needed.	We use this level of computing to do coarse-resolution simulations, which constitute the bulk of the work we do. Increased access to this level of computing would allow us to perform our programmatic work more quickly and easily than we can now. We would use this type of system to perform coarse-resolution simulations of the atmosphere, ocean, and other aspects of the climate system (e.g., carbon cycle). Again, we are ready to perform such simulations now (indeed, are already doing so on other systems). No evaluation of algorithms, etc., would be needed.
26	AtmosChem	Currently, we have 12 papers under development that are requiring simulations. Most would make use of IMPACT in its current form, simply applied in different ways. Each of those simulations could easily use 64-100 processors for weeks to a month.	Such access would allow sensitivity studies of chemical mechanisms, physical parameterizations, input meteorological data, aerosol loading, and other chemistry and physics issues. These studies are essential to our understanding of how to simulation atmospheric chemistry. Also, simply having high quality access to that processor count would allow us to nearly continuously have production runs in the hopper; this is essential to the production of science papers and analysis.
		Access to dedicated time at the 2-8 Tflop level would enable long term "trend" runs that are currently not possible. By this I mean, currently we can do comprehensive chemistry and physics over relatively short time periods (where emissions and activities are held constant), but we can not do the needed long term simulations where trends in emissions are allowed to be integrated into the model simulation allowing one to attempt the simulation of actual multi-decades of atmospheric observations. Only then can one examine the interplay of energy use emissions, infrequent volcanic eruptions, and related activities and how those interact to produce the historical species distributions. This would be a major simulation result; basically the first ever using a model with comprehensive chemistry and physics. It would be a benchmark simulation in the atmospheric chemistry community. Historical emission databases would need to be created for the trend run.	
27	Earthquake	Implicit solution of large coupled fluid-structure system.	

Appendix E—Program Impact

The following table provides a high-level summary of the impact that these projects have on the programs they support.

The column headings used in the table are:

- Project ID—an arbitrary code phrase assigned to this project to allow for quick reference.
- Programmatic Impact of Project—the major program impact or impacts are listed for each project. A one-line **significant impact** for each project appears in bold.

	Project ID	Programmatic Impact of Project
1	ALPS	**Predictive laser plasma interaction (LPI) is fundamental for the design and analysis of laser driven fusion experiments, such as those to be performed at the National Ignition Facility (NIF).** The ALPS (Adaptive Laser Plasma Simulator) project explores the use of parallel adaptive mesh refinement (AMR) in the simulation of LPI. Through the use of AMR, problems with a wide range of scales can be more efficiently solved, which in turn enables larger LPI simulations to be performed in less time. In addition to developing advanced algorithmic and code technologies, we are also involved in testing these new capabilities on LPI problems of importance to the Laboratory's laser-driven high energy density science missions.
		ALPS will be used to model crossed laser beams in a plasma flow, to provide guidance for the experimental program being conducted at the Omega facility at the University of Rochester in support of the NIF program. ALPS will be used to design experiments by helping to decide which experiments to perform, reduce the number of experiments needed to study a large parameter space of plasma conditions and beam conditions. Experimental information will also be used to validate ALPS.
2	DJEHUTY	**Improved determination of the size, age and composition of the universe through 3D modeling of stars.** The LLNL code, Djehuty is now the world's only operating code capable of modeling complete stars in three dimensions (3D). It has an equation of state (EOS) that is very accurate for a large range of masses (>0.5 solar masses), Opal opacities (Rogers and Iglesias 1992) with Alexander opacities (Alexander 1994) for the lower temperatures where molecules are significant, and nuclear reaction-network for hydrogen, helium, and carbon burning. Nuclear energy production can be computed either by a tabular set of reactions, or by a very quick analytic network. Energy transport is modeled by a standard pair of coupled diffusion equations. The gravity implementation is currently complete only for spherical stars, but is adequate to begin the first major 3D study of the convective cores in massive stars. We expect to begin testing of a Poisson solver that will enable us to study rapidly rotating and binary stars. Then, we will enter completely new territory.
3	AMRh	**The accurate calculation of compressible, high-Reynolds-number flows is critical to understand turbulence and mix.** AMRh provides unprecedented spatial and temporal evolution of shock-driven fluid flows and radiatively-driven plasma flows where all the flow length-scales are resolved to the highest degree possible.
4	Fermion MC	**This project is fundamental in understanding opacities and equations of state from first principles.** Scientifically, our ability to calculate the energy and other properties of many-fermion systems is the basic step in developing theories of more complex systems in physics, chemistry, and biology. This project is truly ab initio, requiring only fundamental constants. It serves as a check on experiments, and will lead to replacement of some experiment by computational methods.
5	DD-ICF	**Development of a predictive capability to simulate direct drive Inertial Confinement Fusion** (ICF) implosions by modifying LLNL codes that are currently used for simulating indirect drive implosions. The codes that will be used are HYDRA, a 3D radiation-hydrodynamics code, and Cretin, a 3D code which is used to model atomic line emission from ICF capsules. Developing this simulation capability will provide us a tool for better understanding the relative merits of indirect and direct drive. The exercise of modeling direct drive experiments will also help to validate the codes.
6	Z3	**Development of a state-of-the-art, first-principles predictive capability for LPI (laser plasma interactions) is critically important to the ICF Program, as well as for other programs at the Laboratory.**
		A major application has been the control and mitigation of instabilities that arise in the laser-plasma interactions encountered in the ICF program.
		More recently, these codes have been used to study charged particle production and laser propagation issues relevant to the fast ignitor effort.
		This work was part of the overall experimental and theoretical campaign on the PetaWatt laser.
		Our PIC modeling itself led to an American Physical Society invited talk and a subsequent publication. Our work was also cited in other invited talks and publications. The interesting and intriguing legacy code results form the basis of interest in the next generation PIC codes.
		We used our legacy codes to study the saturation of the Brillouin instability, an important issue in the plasma physics that is expected to occur in NIF generated plasmas for inertial confinement fusion. That work resulted in several publications.
		We currently use the tera cluster for simulations with ZOHAR and BZOHAR both for short-pulse, high intensity studies and modeling relevant to NIF plasmas.

	Project ID	Programmatic Impact of Project
		The renewed interest in modern PIC codes is also a recruiting tool. This past summer, two graduate students (one from Princeton, one from the University of Michigan) spent the summer here developing enhancements to traditional PIC codes for their research.
7	Mat-Shock	**Gaining a basic understanding of interaction of shocks with material microstructure is important to many programs at the Laboratory.** Results to be used as input to continuum models, in particular in relation to anisotropies generated during shock propagation, void nucleation mechanisms and interactions of shocks with defects (dislocations, loops, grain boundaries, etc.).
		Modifications of complex systems such as fused silica due to shocks can now be studied at the atomistic level. These systems are of interest for optics used in high-power lasers such as NIF.
		Using the unclassified computers we have been able to study the nucleation of voids in copper due to propagation of spherical shock waves of interest to DNT and the weapons program.
8	Mat-Rad	**This project's goal is to understand and predict changes in microstructure and properties (mechanical, thermal, dimensional stability) of materials exposed to irradiation**. This is an important problem for many areas. It is crucial for predicting safe operating lifetime of structural materials in nuclear energy technologies (both fission and fusion) as well as for predicting the long-term disposition of materials that undergo radioactive a decay (Pu in stockpile, waste). As shown in Figure 1 the particular conditions of damage rate and defect production can differ by orders of magnitude depending on the specific problem. However the underlying processes of damage production are similar.
		The ultimate objective is to design new, radiation resistant structural materials for fusion energy and predict lifetimes of radioactive materials. Physically-based models are necessary to extrapolate test results to in-service conditions. For example, for fusion environments there is no neutron source available at the moment that operates under the conditions expected during operation.
		We have developed and implemented new interatomic potentials in our parallel molecular dynamics code. In particular we have implemented an empirical interatomic potential for Pu developed at Los Alamos by M. Baskes. Using this potential we have looked at damage evolution in Pu: defect mobility and defect production. These are milestones for our ASCI effort on Pu aging.
		Our work has been presented at the Science Day, UC National Security Review and over 10 International Conferences . It has produced over 20 papers, including Nature and PRB during 2000 and 2001.
9	Cell Modeling	**Understanding the functioning of microbial pathogens at the cellular level is critical to the Lab's biosecurity mission.**
		By manipulating the metabolic, signaling, and regulatory pathways of the microbial cell, we will discover novel ways to shut off bacterial toxin production, defeat genetically engineered organisms, develop new antimicrobial drugs, and develop more robust diagnostic sensors for detecting bioterrorist attacks.
		Understanding the signaling and metabolic pathways of eukaryotic cells and their intercellular communication in tissues will help us to develop better targeting agents for fighting cancer, develop better drug delivery systems, and counter pathogenic attacks on the body.
		Modeling of complex chemical processes via a material-based compartmentalization - modular computational objects of varying complexity. The objects range in complexity from biological chemical species, molecules (amino acids or hormones), proteins which have state, a collection of proteins working as a unit (a protein machine), an organelle, or even a complete cell.
10	FP-Biochem	**Molecular targeting, a technique to engineer antibodies to detect and ultimately kill toxic cells, is an important research area for national security and for medicine resulting from this project.** An important aspect of engineering the antibody is to design a chelator that will bind to a specified metal ion. The demise of the toxic cell is mediated in part by its interaction with this bound metal species.
		Extracting toxic metals such as beryllium from the body is another aspect of receptor-ligand binding is an active area of research at LLNL supported by this project.
		Because molecular level modeling in the area of ligand-receptor binding is at its infancy, a opportunity is presented to set precedence in this area using state-of-the-art computing platforms and algorithms that are present at LLNL.

	Project ID	Programmatic Impact of Project
11	CompBio	**The goal of the computational biology project is to combine LLNL's expertise in biology, advanced simulation methods and high performance computing to develop a laboratory core competency in computational biology.** The primary focus of this project involves the use of state-of-the-art first principles molecular dynamics (FPMD) simulations to examine biologically relevant systems. The use of first principles molecular dynamics enables very accurate dynamical descriptions of biological phenomena including chemical mechanisms, enzyme catalyzed reactions, protein-protein interactions, and DNA-protein interactions.
		This project will move FPMD simulations into the realm of solving a realistic biological system and would be the largest such simulation ever performed. One of the main scientific focuses of the BBRP is DNA repair. In particular one enzyme, Ape1, repairs abasic sites (sites missing nucleotide bases) that can spontaneously arise in DNA and are caused by radiation damage. Ape1 recognized the abasic site and cleaves the DNA backbone (phosphate hydrolysis) such that another enzyme can remove the abasic site and replace it with a "good" DNA base. Currently, solely due to comoputational limitations, we can only represent the DNA as a simple molecule, dimethyl phosphate, with no sugar or base. Also, we cannot even begin to create a model enzyme active site. Given a platform of between 2-8 teraFLOP/s for one to four months, we would be able to create a model enzyme active site and include a small piece of DNA to be cleaved.
		Access to an intermediate capability platform would make it possible for us to examine the differences between a number of variations in an anticancer drug, or to sample reaction coordinates in a model system such as phosphate hydrolysis. In each case, a reaction pathway of a proposed mechanism is simulated. Along this pathway, 10-20 points are chosen to run 3 ps simulations each (10-20 independent simulations) in order to determine the free energy profile of the reaction pathway. Thus, the free energy of activation can be calculated. Each reaction mechanism can be studied and compared to other reaction mechanisms.
		For our current projects, we would like to investigate many of the enzymatic mechanism questions that can only be answered within a first principles molecular dynamics model.
		We would like to increase the size of our simulations to be more biologically relevant (500 atoms and up). To achieve these goals, we are currently limited primarily by computer time.
12	GFMD	**Quantum-based atomistic simulations of materials properties in transition metals is important to the LLNL program on multiscale modeling of strength and failure. This project will allow direct comparison of simulation material-strength simulations with experiment, and provide predictive simulations for extreme conditions of temperature and pressure that are not now accessible to experiment**
		Simulate a full-scale 3D modeling of dislocation junctions that have been observed in TEM experiment. This type of capacity will move our simulation capability as the world-class leader in the field of atomistic simulations of dislocation-dislocation interactions and multi-scale modeling of strength of materials.
13	BOUT Proj	**This project will enable self-consistent modeling of plasma and neutral particle transport in the edge plasma of magnetic fusion devices.**
		The unique BOUT simulation capabilities have led to fundamental discoveries and worldwide recognition. BOUT Simulation results has been used for benchmark with experimental measurements for boundary plasma turbulence research on various magnetic fusion devices, such as DIII-D at GA, C-mod at MIT, and NSTX at PPPL.
		BOUT will be used to uncover the basic physical mechanisms of important edge phenomena and to predict edge plasma behavior for evaluation and optimization of future devices.
		The code will be applied to key issues including the role of non-diffusive and/or large-event-dominated transport, the transition to the enhanced high-confinement mode, edge-localized modes, core density-limit phenomena, and characterization of the loss and fueling channels through the edge plasma.
14	NuclStruct	**The goal of the ab initio nuclear structure project is to attempt a first-principles description of the structure of light nuclei.** In particular, we wish to determine if our knowledge of the fundamental interactions between pairs of nucleons is sufficient to describe the rich and complex structure observed in nuclei. This is a topic of fundamental importance in nuclear physics and our work will represent a significant improvement in our understanding of nuclei.
		These first-principles descriptions of nuclear structure will enhance our ability to accurately calculate cross sections for nuclear reactions. This is important for the nuclear data effort for SBSS within PAT.

	Project ID	Programmatic Impact of Project
15	JEEP	**The main goals of this project are: 1) Predict physical and chemical properties of matter with great accuracy, using advanced quantum simulation techniques, e.g. state-of-the-art first principles molecular dynamics (FPMD) codes. 2) Investigate properties of condensed systems (e.g., fluids and solids), which are not directly accessible to experiments. 3) Interpret and complement experiment in close connections with experimentalists, by taking advantage of high-performance computing. 4) Establish LLNL as a strong player in the field of FPMD simulations worldwide.**
		The combination of expertise in high-performance software development, a growing group of expert code users in various LLNL directorates, and the availability of large supercomputers will potentially make the Laboratory a unique place where the most ambitious molecular simulations can be run. High Pressure Fluids (HPF) -- Microscopic simulations of shock propagation in high explosive mixtures (at 2-8 TF). Microscopic simulations of shock propagation in heavy metals (at 25-100 GF). These two sub-projects could have huge programmatic impact for DNT (both A and B division) and for NIF.
		Semiconductor Nanostructures (SN) -- Optical properties of dots in solution on the fly (at 2-8 TF). Virtual (real time) atomic manipulation of semiconductor nanostructures in solution (at 25-100 GF). These sub-projects could have a huge scientific visibility and eventually will lead to close collaborations with NAI.
		Computational Biology (CB) -- DNA base pairs in solution (at 2-8 TF). Drug attack to DNA in solution (at 25-100 GF). This sub-project would have as well a huge scientific visibility.
		In the last two years, the success of our projects- all of them based on high performance computing- allowed us to hire 11 new scientists.
		Some of the projects initially funded by LDRD are now funded by ASCI because of the programmatic relevance of the results produced.
		In the last two years the jointed JEEP/QSG project has performed two DNT award-winning investigations (1999 and 2001) and a DOE defense programs award of excellence (2000).
		Project has had 32 refereed papers 1999-2001. Members of the project have had about 45 invited talks to international and conferences.
16	PHENIX/HBT	Our scientific goal is to measure and understand the particle emission region for relativistic heavy-ion collisions (i.e. collisions of Au nuclei at light speed). **The relativistic heavy ion program is motivated by the desire to detect and characterize the QCD phase transition (the melting of protons and neutrons into a plasma of their constituent quarks and gluons).** Measuring the particle emission region is an integral part of the search for this phase transition, as well as an attempt to understand the nature of the strong nuclear force during conditions that prevailed in the first microsecond after the big bang.
		We will remove many of the systematic errors made in the traditional data analyses attempting to detect and characterize the QCD phase transition, surpassing what can be achieved from traditional analyses. For the source imaging code, we should be able to answer questions pertaining to non-gaussian source shapes, non-spherical structures in the resonance contribution to the source (the most extended region of the source), and detector effects obtained from the monte carlo simulations can be folded into the source inversion technique.
		We will provide important constraints on current hydrodynamic models of heavy-ion collisions. We will do this through the ability to describe the pion source and expansion with a single parameterization. With additional code development, we should also be able to simultaneously fit multiple species of particles with a single parameterization.
		The group has been directly involved in recruitment of four physicists during the last two years. While continuing to do basic research in the group, most group members have make significant contributions to LLNL programs in areas of advanced radiography, and nuclear reaction modeling.
		This group has been directly involved in recruitment of four physicists during the last two years. While continuing to do basic research in the group, most group members have make significant contributions to LLNL programs in areas of advanced radiography, and nuclear reaction modeling.
17	MD3D	**The scope of the Microscopic Origins of Dynamic Fracture Project is to model, through direct numerical simulation (molecular dynamics), the nucleation and growth of voids in ductile metals during dynamic fracture. The ultimate goal is a model of these processes suitable for continuum hydrocode simulations.** This project represents the first time in which the dislocation mechanisms by which voids grow have been quantified. The development of constitutive models that are sensitive to material microstructure will enable the assessment of changes due to aging and remanufacture.

	Project ID	Programmatic Impact of Project
18	pF3d	**The pF3d code is one of the principal production tools for the ICF Hohlraum Energetics Work Breakdown Statement (WBS-1), part of the theoretical component to the NIF Ignition Plan.** The ability to model laser-plasma interactions (LPI), and ultimately to develop a predictive capability for LPI, is essential in modeling the energetics within a NIF ignition hohlraum, which is key to ensuring successful ignition on NIF. The code pF3d forms the basis for a laser-plasma interaction predictive capability.
		We would compare the results of these full beam simulations to the smaller corresponding letterbox calculations and determine just how well, and when, we can extrapolate. These "capability" calculations on 2-8 Tflop/s system would allow us to make more meaningful use of "capacity" resources.
		Electron thermal conduction must be better understood. By its nature, electron heat conduction in a laser driven hohlraum is nonlocal and nonlinear, making a computational model difficult. This can be studied in 2d using the 25-100GF/s system.
		Saturation mechanisms for both stimulated Brillouin backscatter (SBS) and stimulated Raman backscatter (SRS) can be studied in 2d using the 25-100GF/s system. SBS is generated by the interaction of the laser pump with an ion acoustic wave, while SRS is generated by the interaction with an electron plasma (Langmuir) wave. Part of understanding these saturation mechanisms involves parameter studies with pF3d, in which we vary the frequency of the light, the power in the beam, beam models (with and without smoothing techniques), and plasma conditions.
		We can directly compare pF3d simulation results with results from the particle-in-cell code ZOHAR (fully kinetic, relativistic electromagnetic in 2½d) -- or its relative Bzohar (mobile ions, but Boltzmann fluid electrons), or its daughter code Z3 (3d MPP fully kinetic code) -- for small problems (e.g., a single speckle) to validate the saturation models. Then larger (2-8 Tflop/s) pF3d simulations would be used to study the impact of those models on longer plasmas with RPP beams (when many speckles are present).
19	NIF gas	**Detailed time and space resolved simulations of the flow-field in trapped gas volumes through which the laser propagates are necessary to fully quantify the shot-rate capability of the facility. Fully quantifying the shot rate capacity and identifying ways to improve the shot-rate will have significant impact on NIF operations.** One of the shot-rate limiting factors in the National Ignition Facility will be the recovery of thermally driven gas distortions that result from temperature differences within the laser chain.
		To meet NIF's modeling needs for the laser commissioning effort, these calculations need to be performed in much shorter clock times. To address these needs we are transitioning to parallel tools and platforms. Currently, these calculations are run on high-end workstations and take from several weeks to months for a single simulation.
20	EIGER	**Accurate and timely modeling of electromagnetic phenomena is important to many DOE programs in areas diverse as high energy accelerators, remote sensing and non-destructive evaluation, EMI/EMC effects, lasers and photonics, etc. We are applying our expertise in computational electromagnetics towards the critical problem of electromagnetics effects in high-frequency integrated circuits (IC).** At low frequencies, IC's can be adequately modeled and designed using traditional ciruit theory, which neglets distributed electromagnetics effects. At high frequencies, time-varying electromagnetic fields are distributed throughout the entire circuit, effectively coupling every component of the circuit to every other. This coupling of components via electromagnetic fields is already an important factor in RF and mircrowave IC's, and is a potential show-stopper in the development of next-generation digital IC's with clock rates approaching 100GHz. A computational study of electromagnetics effects within IC's would lead to physics insight and design rules that would positively impact the entire
21	NDE	**The significance of the work performed by this project for a number of programs varies but the common thread is our emerging ability to nondestructively characterize objects and materials at resolution and detail that has not previously been possible.** We are developing nondestructive characterization systems and techniques that require significant computational power. These needs exist due to the complexity of problems that we are dealing with and the size of data sets that are generated or acquired. One of the most significant and evolving challenges that we are facing is processing and visualizing data from large sensor arrays. We are acquiring information at very high-spatial resolution over large fields of view producing extremely large amounts of data.

	Project ID	Programmatic Impact of Project
		We support many programs and WFO efforts. For example, we support the Enhanced Surveillance Campaign (ESC), which is a campaign that supports the Core Surveillance Program. The objective of ESC is to develop tools, techniques, and models that enable us to provide advanced capability to measure, analyze, calculate, and predict the effects of aging on weapons materials and components and to understand these effects as they impact reliability, safety, and performance of weapons that are aged beyond their originally designed lifetimes. We currently have data sets that support this campaign that we cannot process due to a lack of computational resources. We also provide imaging for NIF, the FAA, the Knolls Atomic Power Laboratory, medical applications, and many other organizations.
22	E3D	**Seismic and acoustic wave propagation in the earth and other material is a fundamental physical phenomenon. As such, the ability to model and characterize seismic energy is critically important to a wide range of existing and future Laboratory projects involving scientific, technical, and defense-related applications.**
		As an example, a consortium of companies representing the oil industry recently approached the national laboratories inquiring about the feasibility of performing several (10 to 100) large scale simulations to address current scientific problems encountered during their search for oil and gas. Several of these companies are already collaborators on DOE-sponsored LLNL research projects. The resource requirements for each simulation are approximately 0.5 to 1 TB of memory and 12-48 hours of compute time on a 1 TFlop system.
23	HP-CFD	We develop high performance Computational Fluid Dynamics (CFD) models for simulating flow and dispersion of hazardous materials over urban areas. **The CFD models we have developed are world class and have been used to support chemical-biological projects that are funded by both DOE and DOD.** The problems of relevance to these agencies typically require calculations involving millions of grid points and the use of massively-parallel computer platforms. The results of these model calculations can be used in emergency planning and response activities.
24	NUFT-C	Water movement, contaminant transport, and chemical reactions between migrating fluids and minerals are common processes in the shallow levels of the Earth's crust. These processes play a direct and controlling role in determining important aspects of many environmental, energy and scientific concerns, including groundwater quality and quantity, subsurface pollutant containment and transport, petroleum basin evolution and petroleum migration, oil reservoir management, and sea water intrusion in coastal environments, to name a few. **The NUFT-C code is designed to rigorously account for the coupled physical and chemical processes that occur as sub-surface water migration takes place.** This code is or has been applied to a number of diverse projects and programmatic areas, including Energy and Environment projects in carbon sequestration and high level nuclear waste disposal (the Yucca Mountain Project), and multiple LDRD projects. It has also been applied to problems in other international nuclear waste disposal efforts, including the Japanese nuclear waste repository program.
		The capability to predict how fluid moves in the crust, and what chemical interactions may occur, is critically important to present and future Laboratory interests. Whether considering the impacts of climate change on soil chemistry and properties, contaminant transport and remediation activities, or carbon sequestration in subsurface environments, the processes modelled by the NUFT-C code must be accurately simulated. These and other related areas will play an important role in future Laboratory efforts in the Environment and Energy arena.
25	HR-GCS	**Our project is to perform simulations of present and future global climates at higher spatial resolution than has ever been used in this type of simulation.** Preliminary analyses of our initial high-resolution simulations of present climate indicate that the fine-scale (~100 km) detail in these simulations agrees well with observations. In addition, these simulations produce superior simulations of present climate even on scales that are resolved by coarse-resolution models (~1000 km). Our initial high-resolution simulations of climate changes for the next 100 years show very different results than coarse-resolution simulations in specific regions (e.g. western US, eastern Canada). Thus our simulations are providing (1) improved simulations of climate change on regional (~1000 km) spatial scales; (2) the first fine-scale (~100 km) predictions of global climate change. In short, our simulations appear to have both increased accuracy and increased detail compared to previous simulations. For these reasons, this work is generating strong impact among climate scientists. The increased accuracy and detail in our simulations als
		Strategically, this work is important to the Division, Directorate, and to LLNL because it takes advantage of our strengths—high-end computing and large-scale simulation. We hope therefore to expand our funding and activities in this area. In addition, our strategy of making our results universally available for others to analyze is consistent with the role DOE likes us to play of performing a service for the climate-research community.

	Project ID	Programmatic Impact of Project
26	AtmosChem	T**he goal of our work is to understand the effects of natural and anthropogenic activities on the distribution of important atmospheric chemical species.** These activities include natural events, such as volcanoes, biogenic surface emissions, lightning as well as many others and anthropogenic activities such as energy use, solvents, biomass burning and many others. Implementation of a National Energy Policy will strongly depend on knowing the impact of various energy technologies and policy scenarios on the environment and climate; hence, atmospheric chemistry bridges the issues of energy production and use and environmental quality. Our IMPACT model remains the only model capable of comprehensive tropospheric and stratospheric chemistry, hence, we provide a unique tool in this understanding. We are engaged in studies focused on regulatory analysis (i.e., do we understand the role of emissions of species A and should it be regulated) and also on improved prediction of climate change and its relationship to energy policy.
27	Earthquake	**Dam safety is a very important issue, and if successful, this research will greatly enhance how the U.S. Bureau of Reclamation determines the risk posed to downstream populations.** This research could save the U.S. Bureau of Reclamation dam modification costs.
		This project involves the accurate analysis of seismic, impact, and blast response of various structures. The tools used to accurately simulate the geology, fluid, and structure for this project are useful for many other projects at the Laboratory.

Appendix F—Impact of M&IC

The following table provides a high-level summary of the impact that access to M&IC computing resources has had on these projects in the past.

The column headings used in the table are:

- Project ID—an arbitrary code phrase assigned to this project to allow for quick reference.
- Impact of M&IC Computing to this Project—text describing the impact to the project and associated programs of M&IC computing resources.

	Project ID	Impact of M&IC Computing to this Project
1	ALPS	The resources provided by the M&IC have been essential for the success of our LDRD-funded project. In particular, the TC2K cluster has allowed us to test the ALPS code in ways that would not have been possible otherwise. The ASCI machines were the only other alternative, but they (in particular, Blue Pacific) had become so heavily utilized that it was difficult to get sufficient turn around on batch jobs to make progress. The computations we performed on the TC2K provided a critical part of our successful presentation to the UC committee that reviewed Computation Directorate research in April of this year.
		One of the goals of CASC is to collaborate with Laboratory programs on computational science projects. The availability of institutional computing resources is very important in lowering the artificial organizational and programmatic barriers to the multidisciplinary teamwork essential in pushing the scientific computing envelope. By pooling our computational assets in this way, the Laboratory can accomplish far more than the sum of its otherwise strictly programmatic efforts.
2	DJEHUTY	
3	AMRh	Unclassified computing is essential to the vitality of the laboratory. It is our connection to the outside world. It is also our vehicle for interaction and earning the respect of the outside world. It keeps our scientists current in the latest advances in our respective fields. It definitely provides benefits to the programs by attracting top talent that can impact programmatic work. It encourages significant scientific advances that (at least for hydrodynamics and radiation-hydrodynamics numerical algorithms) advances programmatic computational capabilities.
4	Fermion MC	High performance computing is at the heart of this project. The project is unclassified and includes collaborators who do not have access to our SCF, and whose efforts are essential to the science. Thus the use of M&IC facilities is essential for scientific progress. This project is important as a purely scientific enterprise and is of substantial importance to the mission of the Laboratory. It is world-class work, likely to have high visibility after some additional progress.
5	DD-ICF	Collaboration with Rochester is impractical on classified computers. We cannot send electronic results from the closed network to our colleagues, and some of them may want to run our codes.
6	Z3	We were heavy users of the M&IC computers in the earliest days of the Compass cluster. This work was part of the overall experimental and theoretical campaign on the PetaWatt laser.
		On the compass cluster, we used our legacy codes to study the saturation of the Brillouin instability, an important issue in the plasma physics that is expected to occur in NIF generated plasmas for inertial confinement fusion. That work resulted in several publications.
7	Mat-Shock	The use of the M&IC resources has been crucial for starting a new program within C&MS related to modeling shock propagation in materials at the atomistic level. This has lead to the involvement of C&MS into three new LDRD projects with DNT and Eng. Using the unclassified computers we have been able to study the nucleation of voids in copper due to propagation of spherical shock wave, modifications of silica glass and other materials.
		We have also studied the modifications produced in fused silica due to shocks. This is of interest for optics used in high-power lasers such as NIF.
8	Mat-Rad	Using the unclassified computers we have been able to study very successfully the effect of single high-energy impacts in different materials, related to both ASCI and Fusion Materials programs. We have studied the defects produced by such events as well as the interaction of dislocations with those defects produced during irradiation.
		Access to M&IC computing has also allowed us to hire a new postdoctoral fellow, as well as promote people in our group from postdoc to Term positions.
9	Cell Modeling	This new project is critically dependent on the unclassified capability computing provided by M&IC.

	Project ID	Impact of M&IC Computing to this Project
10	FP-Biochem	This new project is critically dependent on the unclassified capability computing provided by M&IC.
11	CompBio	Unclassified computing is absolutely essential to the success of the projects performed by the Computational Biology group. Without access to large unclassified computing resources, we would be severely limited in the system sizes that we could study.
		Other institutions (such as NIH) are very interested in this relatively new computational tool as applied to biology. However, as noted in the review of one of our recent grant applications, the reviewer's primary complaint focused on the biological relevance because of the short simulation times and small system sizes that we are currently limited to. These problems can only be resolved with access to more computer time. In addition, the reviewers have doubted the commitment of computer time by the LLNL. This problem can also be resolved by specifying that computational biology projects will have a commitment to dedicated computational resources at the lab.
12	GFMD	The resources provided by the M&IC have been essential for the success of our ASCI DofM and LDRD funded High Pressure SI projects. In particular, the TC2K and Linux clusters have allowed us to test the code and make important contribution to our presentations in various ASCI and LDRD review meetings. The ASCI machines have provided very reliable resources for our computational needs, but they had become so heavily utilized that it was difficult to get sufficient turn around on batch jobs to make significant progress. Because of this, we had to rely on other resources such as the one at DOD's Aeronautical Systems Center-Major Shared Resource Center (ASC-MSRC) to meet our unclassified computing needs. We are hoping that M&IC can enhance its current capacity so that the waiting time in the queuing systems can improve significantly through this proposal.
13	BOUT	The unclassified computing (including M&IC) have been valuable resource for BOUT project. Without the resource from the unclassified computing (including M&IC), it is not possible to develop the BOUT project with unique capabilities in the world. BOUT project has had a banner year, as measured by the number of invited papers presented in the major international meetings (five) or in works, by the list of groups who are collaborating with the member of BOUT project (GA, UCSD, Lodestar, IPP/Germany), and by the clamor of experimentalists at GA, MIT and PPPL(including the Lab director at MIT) for more of the BOUT attention.
14	NuclStruct	The resources provided by the M&IC have been essential for the success of our LDRD-funded project. In particular, the Compass, Tera and LX clusters has allowed us to run the MANYEFF code in ways that would not have been possible otherwise.
		M&IC and ASCI resources have provided the single most important tool for us, and our project would not be possible without it. It has enabled us to achieve a significant improvement in nuclear theory that we will now be able to apply data needs of SBSS. Indeed, these resources at LLNL provide a unique capability that has given us a competitive advantage.
15	JEEP	All of the activity of the JEEP and quantum simulation project has been and is at present unclassified. Without the use of unclassified computing, none of our projects would have been possible.
		In the last two years, the success of our projects- all of them based on high performance computing- allowed us to hire 11 new scientists.
16	PHENIX/HBT	M&IC computing has been an essential ingredient in both building and maintaining Heavy Ion Physics group. The DOE Office of Science grant would not have been possible without access to LC computing for running simulations and analyzing data.
17	MD3D	The availability of a large capacity of computational power spread across many projects has enabled sustained scientific progress. These resources have supported the efforts of many scientists including new hires. Our ASCI project has benefited most from the availability of this capacity, and the kind of projects we have undertaken would not be possible without it. The goals of the projects have been designed to make the most of the available computational resources, and they rely on a continued growth. For example, the computational investigation of more complicated materials will require significantly more computational power. The BCC metals are only now becoming accessible by direct molecular dynamics simulations.

	Project ID	Impact of M&IC Computing to this Project
		The fact that LLNL can muster a virtually unequaled computational capability has added tremendously to the prestige of the lab. In a few cases, this capability has permitted qualitatively new kinds of simulations. These simulations are typically not the final word on a subject, but they open the door to new investigations. In other cases they allow a validation of extrapolations that are implicit in many of the calculations.
18	pF3d	For the past several years, the unclassified computing resources provided by M&IC and its predecessors have been essential in meeting our programmatic goals. PF3d modeling of NOVA experiments (hohlraums and CO2 gasbags) accomplished using M&IC resources have contributed to a number of publications ([2]-[6] to name a few).
		Maintaining a cutting-edge computing environment in the unclassified arena will remain essential to us in continuing to meet our programmatic goals for several reasons. Interaction with experimentalists occurs on the unclassified networks, where their data and analysis reside, reporting physics results to our sponsors and the physics community is greatly facilitated by having our simulation results on the unclassified systems, and collaborations with our off-site scientific colleagues can only occur on the unclassified network.
		Having sufficient available computing cycles on the unclassified network has been, and will continue to be, of enormous value to the ICF plasma physics effort.
19	NIF gas	Since the results of out analysis will be part of the NIF distortion information database, the analysis must be performed on unclassified platforms so that the information can be readily transferred.
20	EIGER	We are a big user of the Compass/Tera cluster. Not only is the hardware impressive, but the set of installed software tools is equally impressive, and the sys-admins and the lc-hotline are a great support. We do all of my development/testing/debugging on Compass/Tera rather than on my desktop.
		We use these machines for all of my LDRD research, and this research has led to a $3.6 million DARPA contract.
		We are also supporting 2 Ph.D. students who use these machines for their research.
21	NDE	We currently have data sets that support core Laboratory programs that we cannot process due to a lack of computational resources. We are running into a limit with our current computational resources and we will no longer be able to support our programs unless we get access to significantly expanded computational resources.
22	E3D	The availability of large institutional computing resources, especially through M&IC, has been a critical and necessary component of a diverse set of wave propagation problems important to many Laboratory programs. In addition, these resources have played an invaluable role in developing new technological thrust areas at the Laboratory. The return on investment has been immeasurable.
		Scientific breakthroughs have been achieved in areas of earthquake hazard analysis, crustal deformation modeling, and oil exploration. Results from these investigations have been presented in peer-reviewed journal articles, invited talks, and in presentations at national conferences. Key deliverables have been met that satisfy ongoing programmatic requirements, such as those imposed by the nuclear nonproliferation efforts of the GNEM program. The E3D code and institutional computing resources were used to analyze a critical design issue of NIF, and were used in a collaborative effort with an external medical institute to investigate the detection of cancer in human tissue. E3D and institutional computing has and is being used in a number of LDRD projects. Scientific results from investigations made possible through M&IC resources have received local and national media coverage.
		The availability of institutional computing resources has and is making it possible to explore new focus areas that have the potential to develop into active and well funded programmatic areas directed at the solution of national and international problems. For example, modeling efforts geared toward the detection and characterization of underground structures would not be possible without the capacity and capability computing resources offered through M&IC.

	Project ID	Impact of M&IC Computing to this Project
23	HP-CFD	The computing resource from M&IC is not only essential for us to be able to meet our programmatic goals and milestones, it also enables us to make model predictions to aid instrumentation setup, thus making field experiments more cost-effective.We have been fortunate in gaining access to the ASCI WHITE and, more recently, FROST as beta users. This access of such world class computing platforms enabled us to make a number of state-of-the-art calculations and has drawn much attention from our potential sponsors. This led to several projects with new funding and also enabled us to hire two young scientists into our CFD team. We were also able to propose new, more ambitious, programmatic milestones that would not have been possible without the assistance from M&IC.
24	NUFT-C	In the area of nuclear waste management, there is a strong inertia to overcome that relates to the mind set of “We can’t afford the time and effort to change horses in mid-stream” (in this case, horses being computational methods, i.e., changing from serial to high performance parallel machines). Nevertheless, our efforts are attracting increasing favorable attention nationally and internationally. It is now recognized, as a result of our efforts, that such simulation capabilities need to be adopted. In the foreseeable future, these simulation tools will be employed by the U.S. Yucca Mountain Project, and the Japanese Nuclear Cycle Development Institute through a bilateral agreement. Increasing interest from regulatory elements is also being expressed (particularly by the Nuclear Regulatory Commission and the French Nuclear Energy Agency). It is anticipated that within the next 12-18 months there will be a significant shift by these entities toward reliance on such computational platforms.
		Currently LLNL is the only institution that has exploited this capability and applied it successfully in the nuclear repository arena. Without the resources the lab has developed in this area, we would not be in the leadership position we currently occupy and intend to hold into the foreseeable future.
25	HR-GCS	Our access to M&IC computing has been directly responsible for much of our recent success. Our strategy for future success (indeed future survival) is closely tied to continued access to M&IC computing. Here’s why. The general field we work in (climate modeling) is one in which universities etc. also participate. Since we are much more expensive than universities, we are forced to work on problems that universities cannot address. (If we compete head to head with them, we lose, because we are 2-3x more expensive and not 2-3x more productive.) In climate modeling, the single most important advantage we have over universities is our access to high-end computing.
		In the early days of TC2K, we used dedicated access to part of the machine to perform the highest-resolution near-equilibrium simulation ever performed with a global ocean model. We also performed the highest resolution simulations of direct injection of CO2 into the ocean (as a means of removing carbon from the atmosphere). We have used Frost to perform the highest resolution simulations ever with a global climate model. All of this work has been widely recognized in the scientific community.
26	AtmosChem	Access and use of the M&IC computational resources has been essential and critical to our success. The IMPACT model now stands as one of the worlds most advanced 3-D atmospheric chemistry models. We remain the only location able to carry out combined stratospheric and tropospheric chemistry simulations. Other locations (NCAR and NASA) are beginning the work in this area but are at least 2-3 years away from that capability. These successes come because of the quality staff in ASD and also the state-of-the-art computational resources available to that staff at M&IC. We are the NASA Core modeling site for 3-D chemistry simulations because of our expertise in atmospheric chemistry/science and because of our ability to use large computers. We have been pushing the envelope in coupled chemistry climate modeling over the last few years, much of this done via LDRD and M&IC access. This has resulted in LLNL being selected as the primary institution for implementing interactive chemistry into the NCAR Climate model that is the basis of a large SciDAC project bringing together the skills of NCAR, DOE (LLNL, LANL, ANL,
		Atmospheric chemistry is extremely CPU intensive. Many groups have had to cut back on modeling, run at reduced chemistry and physics, and other methods to get simulation throughput. Access to M&IC have allowed us to move forward with quality models and carry out innovative and first ever simulations.

	Project ID	Impact of M&IC Computing to this Project
		Needs: The large calculations we carry out involve many processors over long periods of time. These sort of capability simulations require high levels of availability and access. There is always a trade off in many users and high access. Many users typically point to short batch queues and long periods of waiting in the queue for simulations to start. This is a highly efficient use of machines and in many situations, it represents optimal use. However, for large capability runs this is not feasible. Over the past years, M&IC has wisely used large machines for relatively few users, thereby providing high quality access. We would like to see that continue. Indeed, simulations obtained via that mechanism have been used in Computations External Review presentations as well as presentations to Vic Reis and other high level DOE officials. These simulations are also those that bring fame and (hopefully) fortune.
27	Earthquake	M&IC has been critical to enabling this project.

Appendix G—An M&IC Capability Resource (MCR)

Background

Linux solutions for high-performance technical computing (HPTC) are now commonplace. Everyone from LLNL and NSF to the American Museum of Natural History in New York City is building Linux clusters based on commodity parts (nodes, interconnect) and Open Source software (primarily Linux).[1] A listing of the TOP500[2] shows over 100 clusters for HPTC over 30 GF/s. The NSF Distributed TeraScale Facility (DTF)[3] Linux cluster (11.6 TF/s in mid 2002 for $53M) is actually physically distributed over four sites (NCSA, ANL, SDSC, and CalTech).

Our recent experience with Linux Parallel Capacity Resource (PCR) clusters for the NNSA ASCI ongoing computing element indicate three very important factors:

- Performance of commodity (Intel, IA-32) clusters is comparable, if not faster, on LLNL applications to proprietary RISC solutions.
- The price of these clusters is less than proprietary RISC solutions by a factor of 3 to 10.
- Linux-based clusters now scale from tens to thousands of nodes.[4]
- Support for Linux clusters can be more challenging than that for proprietary Unix solutions, but as we are discovering with our existing SSP Linux system, this is a tractable problem.

ASCI ongoing computing paid approximately $2.7M for 1.5 TF/s in three clusters (129, 89, and 29 nodes each) in FY01. These clusters were delivered in September 2001 and are now producing scientific results in "science runs" mode. We anticipate migrating the 89-node cluster to classified operation in early February 2002 and the 129-node cluster in March 2002. These machines are targeted at capacity MPI jobs (4–32 MPI tasks) that are not appropriate for the large ASCI platforms (ASCI Blue-Pacific and White) and too large for current desktops or Compaq capacity clusters.

On the basis of the requirements described in detail in this White Paper and utilizing the ASCI Linux cluster experience, we will propose to scale up this cluster architecture to much higher node counts for a capability platform for M&IC.

Building Blocks

The PCR clusters are based on Dual IA-32 nodes with two 1.7-GHz Pentium4 processors, 2.0 GB of RDRAM memory, and 72 GB UDMA66 disk drives. The peak of the nodes is 3.4 GF/s, delivering about two times the performance of TC2K per CPU, two to four times the

---

[1] See http://www.amnh.org/science/genomics/research/computing.html and http://www.amnh.org/programs/conference/supercomputing/ for more information.

[2] http://clusters.top500.org/db/Query.php3

[3] See http://www.nsf.gov/search97cgi/vtopic for more information on the NSF DTF.

[4] http://www.ncsa.uiuc.edu/UserInfo/Resources/Hardware/IA32LinuxCluster/TechSummary/

performance of White, and four to six times the performance of Blue-Pacific. The interconnect is Quadrics QsNet, delivering 4.5 ms latency and 240 MB/s unidirectional and 360 MB/s bidirectional bandwidth to user applications. This is the fastest commodity interconnect available. In addition, the cluster is based on Linux OS and Open Source cluster tools and NFS global file system. The code development environment is based on Open Source GNU compilers, the proprietary Intel compilers, and proprietary TotalView debugger. These tools, except for the Intel, KAI, and PGI compilers, are available on all of the other LC platforms. For the M&IC Capability Resource (MCR), we propose to add the Lustre global file system[5] technology.

M&IC Capability Resource (MCR) Hardware

We propose to build a state-of-the-art Linux cluster (Fig. G.1) that will provide between 80% and 200% of the delivered performance of the world's fastest ultracomputer (ASCI White) on M&IC scientific applications at a cost of more than a factor of 6 lower than that of ASCI White! This is a cost performance advantage of about 5 to 12 times for the proposed MCR cluster. For instance, the projected Linpack performance on this machine is 5.55 TF/s. If placed on the current TOP500 list[6] this machine would rank second in the world behind White.

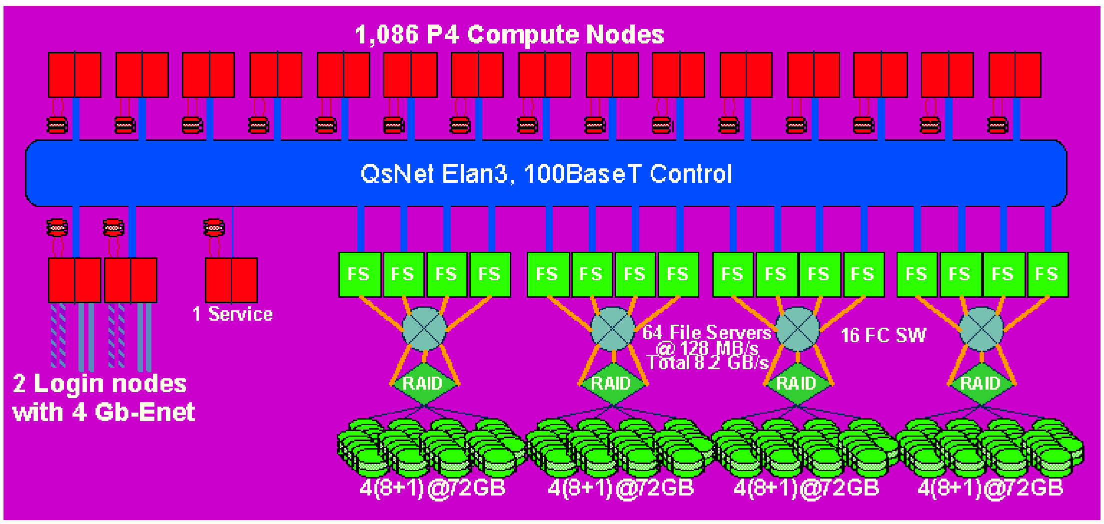

Figure G.1. 11.1-TF/s MCR architecture has 1086 compute nodes that can support multiple simultaneous jobs up to 2172 MPI tasks. Jobs requiring powers of 2 can be accommodated up to 2048 MPI tasks. It also supports a global file system with 8.2 GB/s delivered I/O performance.

The proposed MCR cluster is 11.1 TF/s peak, 2.3 TB of memory, 92 TB of local disk, and 148 TB of global disk aggregate capability with 1152 Dual Pentium4 nodes with 2.4 GHz Pentium4 Xeon processors, 2.0 GB of local memory, and 80 GB local disk.

---

[5] http://www.lustre.org/

[6] Eighteenth list published Nov. 10, 2001, http://www.top500.org/list/2001/11/

The proposed network is Quadrics ELAN3 federated switch. When the PCR was purchased, the ELAN3 network was limited to a single 128-way switch. For the LANL Q and Pittsburg Supercomputer Center (PSC) machines (and other Compaq Sierra systems), Quadrics extended the ELAN3 to a multistage network. For the MCR cluster we propose to utilize 16 ELAN3 128-way switches in the following configuration (see Fig. G.2). Twelve of the switches would have 96 (Down) ports each connected to MCR nodes ($96 \times 12 = 1152$) the remaining 32 (Up) ports would be utilized to connect to the four second-level 128-way switches. This provides excellent connectivity for M&IC applications. Three key figures of merit for MPI interconnects are latency (this determines the short MPI message performance), link bandwidth (long MPI message performance from a single node) and bisection bandwidth (long MPI message performance across the entire machine). For this configuration, the latency is below 5.0 µs (compared with 17–85 µs for White). For the link bandwidth, the key figure of merit is how much bandwidth there is as a fraction of the node peak performance (B:F ratio). For this configuration B = 680 MB/s (counting both directions) and the node peak is 9.6 GF/s. Hence the node B:F ratio is 0.68/9.6 = 0.071 (compared with 0.083 for White). The bisection bandwidth B:F ratio is 130.56 GB/s bandwidth to 11.1 TF/s. So the network bisection B:F ratio is 0.131/11.1 = 0.012 (compared with 0.041 for White). This is an extremely well-balanced and cost-effective network for M&IC applications, as indicated by the performance scaling numbers obtained so far on the PCR cluster.

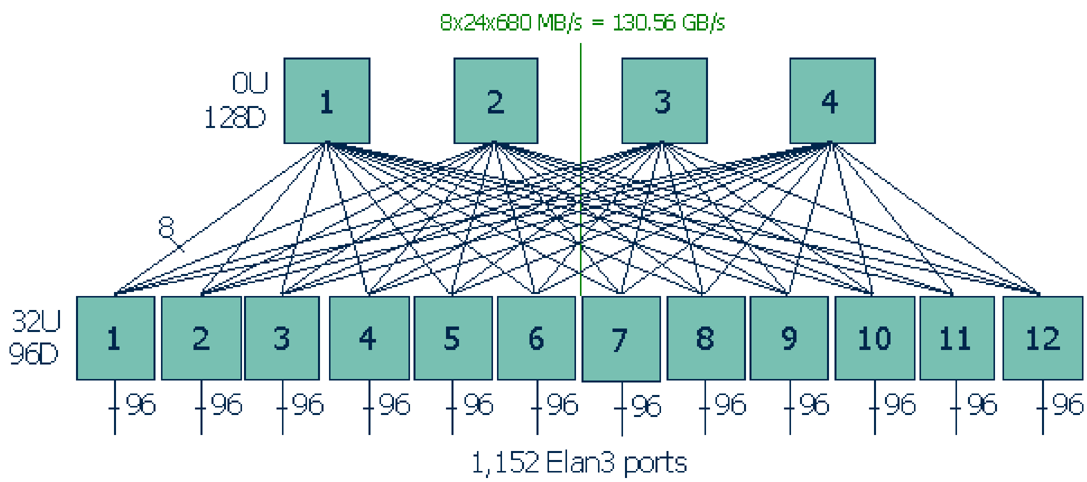

Figure G.2. QsNet ELAN3 network for MCR has a link bandwidth of 680 MB/s (B:F of 0.071) and bisection bandwidth of 131 GB/s (B:F of 0.012).

Hardware Options

The MCR cluster hardware outlined above is highly scalable. There are a significant number of natural scaling points based on the federated QsNet ELAN3 interconnect. For instance, half size (with six level-one 128-way switches and two level-two 128-way switches) is a natural sizing. The final size of the cluster will be driven by budget considerations and by these technical scaling issues. In addition, there will probably be some flexibility in the I/O configuration depending on who wins competitive procurement. We expect at least the following bidders: HP, IBM, Dell, Linux NetworX, Compaq, and Cray.

The previous discussion assumed Dual 2.4 GHz Foster based Pentium4 nodes with the Intel i860 chipset and RDRAM and motherboards that became available last summer. These motherboards are now well qualified, have great delivered memory bandwidth performance, have LinuxBIOS implementations, and have the highest GHz rating available. However, the next-generation motherboards based on Intel's next generation Pentium4 implementation (code named Prestonia) and associated Plumas chipsets are now sampling and will be available in volume in March 2002. The advantages of this option are (1) double the L2 cache size to 512 KB, (2) SDRAM memory interfaces, (3) improved PCI performance for Quadrics, and (4) 1U form factor (half that of the current PCR nodes). This means that the interconnect bandwidth of these nodes would go from around 220 MB/s to 300–310 MB/s. In addition, the SDRAM memory would be much cheaper (by a factor of 2 or 3). However, the drawback to this approach is that Prestonia will probably be available at 2.0 to 2.2 GHz, and we are unsure of the actual delivered memory bandwidth. We suggest that the ICEG run benchmarks on the new nodes and determine if they are attractive.

M&IC Capability Resource (MCR) Software

The software environment for MCR is based on three principles: (1) keep the code development and production environment as similar as possible to those of other cluster resources (ASCI and M&IC clusters); (2) build as much of the environment as possible from Open Source; (3) provide a complete, stable "production" environment.

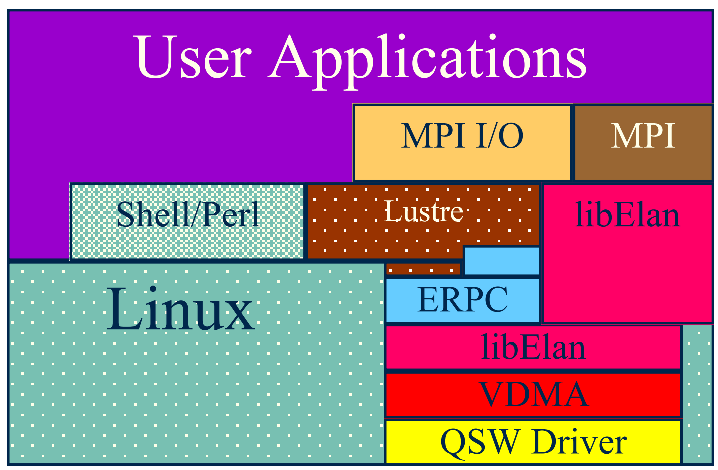

Figure G.3. The MCR software stack is based primarily on Open Source components. It provides a complete code development and production environment.

Except for changing the processor architecture from Compaq Alpha (TC2000 and capacity clusters) or IBM Power (ASCI White) RISC to Intel IA-32, the rest of the hardware environment for MCR is very similar to that of other large computing clusters at LLNL. This hardware strategy eases, but does not completely eliminate, the chore of application migration to commodity hardware and Open Source software based systems. For instance, the key programming environment characteristics remain: (1) each node has a Linux OS with local disk; (2) global I/O through the Lustre file system is provided through a client server model; (3) users log in to specialized nodes that have access to the global file system and provide high-speed parallel FTP access off the machine. In addition, the full suite of code development tools is available: (1) Intel, PGI, and GNU compilers; (2) TotalView debugger; (3) hardware performance and MPI tracing tools for performance analysis. Fig. G.3 shows the software stack from the perspective of user applications development.

The production software environment provides the same tools to run batch and interactive jobs as do the other clusters at LLNL. Fig. G.4 shows the interaction between the user and system services to provide interactive job launch.

Except for compiler bug issues, code porting to MCR should be fairly straightforward.

MCR Architecture Risk Analysis

The size of this cluster, at 1152 (or the half-sized variant), poses some interesting challenges. Not many Linux systems have been scaled to this level. In addition, there is currently no Open Source global file system available for large clusters (NFS is sufficient for small clusters). These are the two major risk areas for the MCR architecture. The Linux scalability risk has been addressed by enlisting the support of software partners—RedHat for Linux support and Quadrics for the interconnect and job launch software. RedHat[7] is a world leader in providing Linux distributions and enterprise level support. We are putting a contract in place for at least one on-site support person that will do half-time kernel development, debug, and testing and half-time general Linux and tools support. This person will be critical for providing LC with an inside track into the Linux kernel development efforts and RedHat's distribution and support organizations.

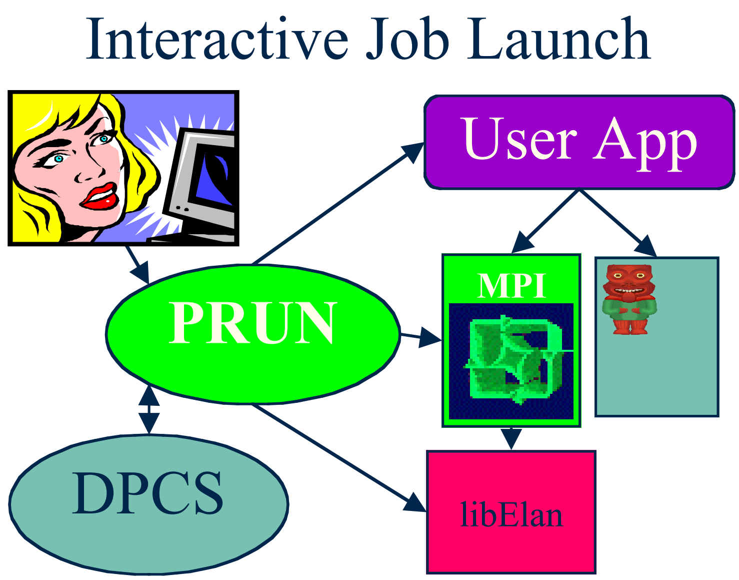

Figure G.4. Interactive (and batch) job execution environment is similar to that of other clusters at LLNL (e.g., TC2K, ASCI White).

For the global file system, we have aligned ourselves with the long-range ASCI strategy. To this end, we have engaged Peter Braam and his technical contributors at Cluster File Systems, Inc. (CFS) to provide us with a simplified implementation of Lustre (called Lustre Lite, of all things) that will be available in the time frame for the MCR deployment. The contract with CFS is similar to the RedHat contract in that we are fostering development with an industry luminary and they will be providing at least one FTE effort directed at MCR. In addition, another governmental laboratory has recently chosen an Intel-based Linux cluster as their next supercomputer, and they will be collaborating with us on this effort. Their vendor partner is a Tier One vendor who will probably be a strong bidder on Purple.

Timeline

One critical thing we learned with the PCR cluster procurement and integration is that it is very important to avoid delivery and acceptance milestones (with associated payments) at fiscal year boundaries. With this in mind, we offer the following tentative schedule (Table G.1), with the

[7] http://www.redhat.com

aim of providing sufficient buffer (2 months) to get the system working before having to pay the vendors or have the money evaporate at the end of the fiscal year. This is a high-level consideration because payment is one thing that really motivates these folks.

Table G.1. Tentative schedule for PCR cluster procurement.

Target Date	Milestone
January 28, 2002	White Paper complete
March 15	ICEG finalize node choice
April 1	Finalize budget and procurement strategy
April 7	ICEG finalize configuration
April 15	Release RFP to industry
May 15	Select vendor
May 21	Complete contract negotiations and award contract
June 1	Start MCR build
July 1	Start MCR pre-ship test
July 7	MCR delivery
July 21	Start MCR acceptance testing
August 1	Start system debug for Science Runs
September 1	Start LLNL Science Runs
January 15, 2003	Start MCR Limited Availability
February 15, 2003	Start MCR General Availability

This schedule is carefully calibrated with the other major activity in the LC this year, the ASCI Purple procurement and Purple Early Deployment of Technology Vehicle (EDTV) delivery. In addition, there are short, but not aggressive schedules for the procurement, build and acceptance, and debugging phases of the project. Still, even with these short time periods, **it is essential that the budget be set (to rough order of magnitude) by April 1, 2002**. Otherwise, we will have end-of-year problems and the project risk increases significantly.

Appendix H—Configurations and Resource Scheduling

The possible addition of a large Linux cluster (MCR) to the institutional computing resources at LLNL will substantially increase the resources available. We will not only be able to service more of the existing work, but will be capable of addressing larger problems than previously possible on TC2K alone. Our analysis of the responses to the Fall 2001 M&IC Questionnaire clearly indicated that the majority of projects require access to both highest-end *capability* computing (~10 TF/s) and *large-capacity* computing (~100 GF/s). Access to both of these classes of computing resources can be accomplished through the flexible partitioning of a single large system or through other strategies. To get the most benefit from these systems, it is very important to configure the systems and their scheduling parameters appropriately. Proposed configuration options are presented below.

Current Capability

The institutional computing requirements are currently served by the Compaq TC2K system. We have a wealth of historic information on that system that should be of some value in forecasting MCR workload characteristics.

TC2K has a 10-node debug partition with limits of four nodes per job, two jobs per user, and a maximum run time of 1 hour. TC2K also has a 118-node batch partition with limits of 32 nodes per job, three jobs per user, and a maximum run time of 12 hours. Overall system utilization runs about 70% of available resources, which is respectable for a parallel system. Average job size is about 14 nodes, with about one quarter of the jobs being the 32-node maximum. While the majority of jobs terminate within 1 hour, the average execution time (weighted by actual resource use) is 7.6 hours. Almost half of the jobs run for the 12-hour system limit. The average queue wait before beginning execution is 72 minutes, although some jobs wait many hours for initiation. The overall picture is that of a heavily utilized parallel computer supporting both a substantial "production" workload and a significant debugging effort.

Proposed Configurations

There are many configuration options for the MCR cluster. Some attractive options are presented below. These options include a reconfiguration of TC2K to provide capabilities complementary to those of the proposed MCR system. Our premise is to support, at least on a regular basis, the execution of much larger jobs than is currently possible. Although the final MCR configuration has not yet been established, the proposals below are based on having 541 compute nodes, each with two processors, to provide a concrete example for exploration. Other configurations, either smaller or larger, could be handled by analogous configurations.

Both TC2K and the proposed MCR system use a Quadrics interconnect. It is critical to understand a limitation of this interconnect in defining machine configurations. Quadrics supports a message broadcast capability, but only if all tasks of the job are allocated to consecutive nodes. Any job allocated nonconsecutive nodes will process message broadcast requests as sequential communications to each node. We have not yet (but will) researched the performance degradation, but there are reports of job execution times in certain cases increasing by 100% or more. Quadrics is working on enhancements to ameliorate the situation, but we do not have a date for availability. Currently, the placement of jobs onto specific nodes is critical, especially for larger jobs.

DPCS considers each node in a cluster as equivalent to every other node. While this paradigm functions well on many computers and simplifies scheduling, it can result in severe inefficiencies on Quadrics interconnect systems. Until the improved Quadrics software is available, the specific node allocation to jobs will be critical, and DPCS is not designed to consider this.

Option: Stand-Alone Configuration

The MCR cluster can be configured in a similar fashion to TC2K, with a relatively small debug partition and large batch partition. Larger batch jobs could be permitted to execute at all times, as opposed to what is currently possible on TC2K. The execution of very large jobs (say over 25% of resources) poses some challenges. System utilization would be severely reduced if such jobs were permitted to execute at all times, because resources would need to be accumulated over the course of hours in order to initiate the larger jobs. From the perspective of system utilization, a more attractive option would be to execute very large jobs only on weekends. If longer execution times were permitted, even this option would cease to be viable without premature termination of running jobs. Large jobs would also routinely fail to be allocated consecutive nodes because of fragmentation and would therefore suffer severe performance degradation.

Option: Stand-Alone Configuration with Partitioning

The MCR cluster can be configured with three partitions: a small debug partition, a normal batch partition, and a large-size job batch partition. DPCS would be configured so the customer could direct jobs to a specific partition. The large-size job batch partition would execute only jobs having the same size as the partition (an option might be to allow somewhat smaller jobs to use this partition, but they would be "charged" for using the entire partition by the fair share scheduler). This partitioning enhancement offers several significant advantages:

- Large-size jobs would be able to execute at all times.
- Large-size jobs would be guaranteed consecutive nodes and thus would avoid the Quadrics interconnect performance degradation described above. This is a very important advantage of this configuration.
- Moderate-size jobs could be permitted longer execution periods without adversely affecting the scheduling of large-size jobs (no resources wasted accumulating resources to start the larger jobs).

The artificial partitioning of the computer could adversely impact overall system utilization if there were an imbalance in workload for the two batch partitions. Assuming a backlog of work for both partitions (something LC and the user community would need to work on together to assure is the case), system utilization should be very good (probably over 80%).

Some possible configuration numbers that might be considered are shown below.

Partition Name	Partition Size	Job Size Limit	Job Time Limit	Notes
Large batch	128 nodes	96–128 nodes	24 hours	Charged for entire partition
Normal batch	393 nodes	96 nodes	12 hours	
Debug	20 nodes	4 nodes	1 hour	Two jobs per user

Note that with the 2.3-GHz processors proposed, 128 nodes translates to about 1.2 TF peak, or a little less than double the peak of TC2K.

Option: With Reconfiguration of TC2K

As already noted, TC2K supports both a significant production computing workload and substantial development work. We could continue to execute smaller node-count jobs on TC2K and direct both the development and highly parallel work primarily to MCR. Jobs requiring more than 32 nodes would only be executed on MCR, and the system could be made responsive by keeping a relatively short time limit for all batch jobs, say 6 hours. With much of the development work shifted to MCR, time limits on TC2K might be increased to 24 hours. Both systems would have small debug partitions. This scenario offers somewhat better economies of scale but removes some customer flexibility. A summary of these proposed limits is shown below. Ownership shares on TC2K would need to be divided between MCR and TC2K if this mixed option were adopted.

TC2K old limits	32 nodes	12 hours	384 node-hours
TC2K new limits	32 nodes	24 hours	764 node-hours
MCR limits (normal batch)	96 nodes	6 hours	1528 node-hours
MCR (large batch)	96–128 nodes	24 hours	3072 node hours charged

Option: Weekend Operation

It would be possible to treat the above options as weekday operations and then to establish a weekend process that would allow production codes to run up to 256 or even 512 nodes for a day or longer.

ICEG Guidance is Required

This subject is technically complicated; technological solutions are limited; and functionality issues are poorly understood by most users, who tend to become impatient with limits established by logic based on experience with which they are unfamiliar. Each of these proposed configurations has some advantages; other configurations are possible. We encourage you to send comments and suggestions to tomaschke1@llnl.gov for consideration by the ICEG. As usual, there is the tension between those who would like to run calculations for very long periods, those who seek rapid access to the system, and LC that seeks (for obvious accounting oversight issues) full utilization of the resource. No solution is even close to perfect. The best we can hope for is a reasonable solution and consensus as to the trade-offs.

Appendix I—The Promise of BlueGene/L

BlueGene/L is the name given to a scaleable ultracomputer that is being developed through a collaboration between IBM and NNSA. The system, based on a new architecture, has a tight coupling between the applications R&D and the hardware and systems R&D to deliver exceptional processing power of 180 TF/s for a wide range of applications. Additionally, the architecture can be leveraged to reach 360 TF/s for applications that require significantly more computation than communication or that require only nearest-neighbor communications. This machine is targeted for delivery in the 2004–2005 time frame, at a price/performance unobtainable with conventional architectures.

BlueGene/L is designed to deliver this exceptional price/performance on a selected set of science applications. While it is not now envisioned as a replacement for current ASCI platforms, it will have a significant impact on the Stockpile Stewardship Program through high-confidence simulation of physical phenomena. If successful, BlueGene/L could foster future ASCI platform contenders, and this approach is expected to scale to capabilities of a PF/s and beyond.

Table I.1. Characteristics of BlueGene/L and those of other ASCI platforms.

	Red	Blue Mountain	Blue-Pacific	White	Q	BlueGene/L
Machine Peak Speed (Tflops/s)	3.15	3.072	3.89	12.3	30	180/360
Total Memory (TBytes)	2.4	1.5	2.6	8	33	16-32
Footprint (sq ft)	2,500	9,072	5,120	10,000	20,000	2,500
Total Power (MW)	0.825	1.500	0.59	1.0	3.8	1.2
Cost (M$)	56	121	97	110	215	~50 est.
Installation Date	8/1998 (1.0872 TF), 9/1999 (3.15 TF)	10/1998	10/1998	9/2000	~6/2002	~12/2004
Node Type	custom node, Intel Xeon CPU	SGI Origin 2000, MIPS 10000 CPU	IBM SP, Silver node, PowerPC 604e CPU	IBM SP, Nighthawk2 node, Power3+ CPU	Compaq ES45 node, Alpha EV68 CPU	IBM PPC440 SOC, dual CPU, quad FPU
# Nodes	4,680	48	1,496	512	4,096	65,536
CPUs Per Node	2	128	4	16	4	2
Clock Frequency (MHz)	333	250	332	375	1000	700
Power Dissipation Per Node (W)	176	31250	394	1953	922	15
Peak Speed Per Node (Gflop/s)	0.67	64	2.6	24	7.3	2.8
Memory Per Node (GiB*)	1	32	3	16	8	0.250
Interconnect Type	Dual Plane Mesh	HIPPI-800	TB3	Colony DS	Quadrics ELAN3	3D torus, global tree, Gb Ethernet (I/O nodes)
MPI Latency (microseconds)	13	150	17	25	4.5	5
Interconnect Link Bandwidth (B:F)	0.6	0.038	0.11	0.083	0.085	1.5
Interconnect Bi-Section Bandwidth (B:F)	0.028	0.019	0.03	0.042	0.043	0.008

The computer architecture being developed will address five key problems facing scalable ultra-computer design and usage: (1) power and cooling; (2) floor space; (3) cost of systems; (4) interconnect cost and performance; (5) distance to memory (single node performance). The design point of BlueGene/L utilizes IBM PowerPC embedded CMOS processors, embedded DRAM, and system-on-a-chip techniques that allow for aggregation of all system functions including compute processor, I/O message processor, three cache levels, and interconnection network onto a single ASIC. This yields a scalable ultracomputer with extremely cost-effective characteristics and extremely low power (~1 MW), cooling and floor space (~2500 sq ft) requirements. These characteristics are an order of magnitude less than the comparable projected requirements for other mainline supercomputer platforms being delivered in this time frame.

Table I.2. The aggressive development and build schedule for BlueGene/L results in a final delivery to LLNL as soon as the fourth quarter of 2004.

Quarter	Event
2Q CY02	ASCI Design Review
4Q CY02	Prototype Design and Cost Review
1Q CY03	Prototype assessment (input to build decision)
3Q CY03	Prototype Design Review
4Q CY03	512-node prototype available at LLNL
3Q CY04	Delivery of first 8K nodes to LLNL
4Q CY04	Final delivery to LLNL

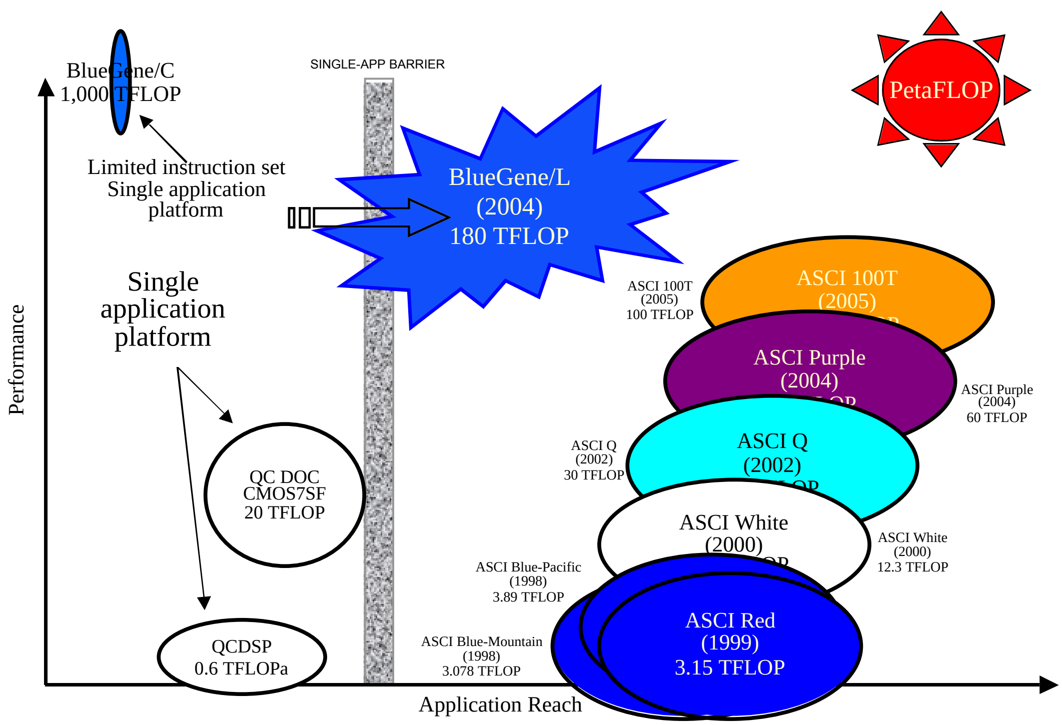

Figure I.1. BlueGene/L is different than the other ASCI platforms and occupies a different spot in the ultracomputing landscape.